

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1912

CCATAACCAT GAGTGATAAC AC

22

10

2) INFORMATION FOR SEQ ID NO: 1913

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1913

CGCCTTGATC ATTGGAACC

20

25

2) INFORMATION FOR SEQ ID NO: 1914

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1914

CGCCTTGATC GTTGAACC

20

40

2) INFORMATION FOR SEQ ID NO: 1915

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1915

CGCCTTGATA GTTGGGAACC

20

5

2) INFORMATION FOR SEQ ID NO: 1916

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 20 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1916

CGTGGGTCTT GCGGTATCAT

20

20

2) INFORMATION FOR SEQ ID NO: 1917

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 20 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1917

CGTGGGTCTG GCGGTATCAT

20

35

2) INFORMATION FOR SEQ ID NO: 1918

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
45 (D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1918

GTGGGTCTCA CGGTATCATT G

21

50

2) INFORMATION FOR SEQ ID NO: 1919

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1919

CGTGGGTCTC TCGGTATCAT T 21

15

2) INFORMATION FOR SEQ ID NO: 1920

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1920

CGTGGNTCTC GCGGTATCAT 20

30

2) INFORMATION FOR SEQ ID NO: 1921

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1921

CGTGGGTCTA GCGGTATCAT T 21

45

2) INFORMATION FOR SEQ ID NO: 1922

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 bases
(B) TYPE: Nucleic acid

50

- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1922

GTTTTCCAAT GATTAGCACT TTTA

24

10

2) INFORMATION FOR SEQ ID NO: 1923

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1923

GTTTTCCAAT GATAAGCACT TTTA

24

25

2) INFORMATION FOR SEQ ID NO: 1924

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1924

GTTTTCCAAT GCTGAGCACT TTT

23

40

2) INFORMATION FOR SEQ ID NO: 1925

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1925

CGTTTTCCAA TGATGAGCAC TTT

23

5

2) INFORMATION FOR SEQ ID NO: 1926

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 23 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1926

GTTTTCCAAT GGTGAGCACT TTT

23

20

2) INFORMATION FOR SEQ ID NO: 1927

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 861 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 35 (A) ORGANISM: *Neisseria meningitidis*
(B) STRAIN: MC9690-129
(C) ACCESSION NUMBER: AF126482

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1927

| | | |
|----|--|-----|
| 40 | ATGAGTATTC AACATTTTCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT | 50 |
| | TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG | 100 |
| | CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC | 150 |
| | AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT | 200 |
| | GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG | 250 |
| 45 | CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG | 300 |
| | GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT | 350 |
| | AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCTGCCA | 400 |
| | ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG | 450 |
| | CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT | 500 |
| 50 | GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA | 550 |
| | TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT | 600 |
| | TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC | 650 |
| | ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG | 700 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GAGCCGGTGA | GCGTGGGTCT | CGCGGTATCA | TTGCAGCACT | GGGGCCAGAT | 750 |
| GGTAAGCCCT | CCCGTATCGT | AGTTATCTAC | ACGACGGGGA | GTCAGGCAAC | 800 |
| TATGGATGAA | CGAAATAGAC | AGATCGCTGA | GATAGGTGCC | TCACTGATTA | 850 |
| AGCATTGGTA | A | | | | 861 |

5

2) INFORMATION FOR SEQ ID NO: 1928

10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: HB251
- (C) ACCESSION NUMBER: X57972

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1928

| | | | | | | |
|----|------------|------------|-------------|------------|------------|-----|
| 25 | ATGAGTATTC | AACATTTCCG | TGTCGCCCTT | ATTCCCTTTT | TTGCGGCATT | 50 |
| | TTGCCTTCCT | GTTTTTGCTC | ACCCAGAAAC | GCTGGTGAAA | GTAAAAGATG | 100 |
| | CTGAAGATCA | GTTGGGTGCA | CGAGTGGGTT | ACATCGAACT | GGATCTCAAC | 150 |
| | AGCGGTAAGA | TCCTTGAGAG | TTTTCGCCCC | GAAGAACGTT | TTCCAATGAT | 200 |
| | GAGCACTTTT | AAAGTTCTGC | TATGTGGCGC | GGTATTATCC | CGTGTTGACG | 250 |
| 30 | CCGGGCAAGA | GCAACTCGGT | CGCCGCATAC | ACTATTCTCA | GAATGACTTG | 300 |
| | GTTAAGTACT | CACCAGTCAC | AGAAAAGCAT | CTTACGGATG | GCATGACAGT | 350 |
| | AAGAGAATTA | TGCAGTGCTG | CCATAACCAT | GAGTGATAAC | ACTGCGGCCA | 400 |
| | ACTTACTTCT | GACAACGATC | GGAGGACCGA | AGGAGCTAAC | CGCTTTTTTG | 450 |
| | CACAACATGG | GGGATCATGT | AACTCGCCTT | GATCATTGGG | AACCGGAGCT | 500 |
| 35 | GAATGAAGCC | ATACCAAACG | ACGAGCGTGA | CACCACGATG | CCTGCAGCAA | 550 |
| | TGGCAACAAC | GTTGCGCAAA | CTATTAAC TG | GCGAACTACT | TACTCTAGCT | 600 |
| | TCCCGGCAAC | AATTAATAGA | CTGGATGGAG | GCGGATAAAG | TTGCAGGACC | 650 |
| | ACTTCTGCGC | TCGGCCCTTC | CGGCTGGCTG | GTTTATTGCT | GATAAATCTG | 700 |
| | GAGCCGGTGA | GCGTGGGTCT | CGCGGTATCA | TTGCAGCACT | GGGGCCAGAT | 750 |
| 40 | GGTAAGCCCT | CCCGTATCGT | AGTTATCTAC | ACGACGGGGA | GTCAGGCAAC | 800 |
| | TATGGATGAA | CGAAATAGAC | AGATCGCTGA | GATAGGTGCC | TCACTGATTA | 850 |
| | AGCATTGGTA | A | | | | 861 |

45

2) INFORMATION FOR SEQ ID NO: 1929

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 808 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella oxytoca*
 5 (B) STRAIN: 26W
 (C) ACCESSION NUMBER: U09188

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1929

```

10 ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
   TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100
   CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAGCT GGATCTCAAC      150
   AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT      200
   GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG      250
15 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG      300
   GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT      350
   AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA      400
   ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG      450
   CACAACATGG GGGATCATGT AACCCGCCCT GATAGTTGGG AACCGGAGCT      500
20 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA      550
   TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT      600
   TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC      650
   ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG      700
   GAGCCGGTAA GCGTGGATCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT      750
25 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC      800
   TATGGATG                                     808
  
```

30 2) INFORMATION FOR SEQ ID NO: 1930

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
 (B) TYPE: Nucleic acid
 35 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

40 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
 (C) ACCESSION NUMBER: AF190692

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1930

```

45 ATGAGTATTC AACATTTTCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
   TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100
   CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC      150
   AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT      200
50 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTTGACG      250
   CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG      300
   GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT      350
   AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCTGCCA      400
  
```

ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG 450
 CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT 500
 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA 550
 TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT 600
 5 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC 650
 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG 700
 GAGCCGGTGA GCGTGGGTCT GCGGTATCA TTGCAGCACT GGGGCCAGAT 750
 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC 800
 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA 850
 10 AGCATTGGTA A 861

2) INFORMATION FOR SEQ ID NO: 1931

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 20 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 25 (A) ORGANISM: *Escherichia coli*
 (B) STRAIN: BM2728
 (C) ACCESSION NUMBER: AF104442

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1931

30

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT 50
 TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG 100
 CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC 150
 AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT 200
 35 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTGTTGACG 250
 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG 300
 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT 350
 AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA 400
 ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG 450
 40 CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT 500
 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA 550
 TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT 600
 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC 650
 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG 700
 45 GAGCCGGTGA GCGTGGGTCT CTCGGTATCA TTGCAGCACT GGGGCCAGAT 750
 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC 800
 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA 850
 AGCATTGGTA A 861

50

2) INFORMATION FOR SEQ ID NO: 1932

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 5 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 10 (A) ORGANISM: *Escherichia coli*
 (C) ACCESSION NUMBER: AF190695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1932

```

15 ATGAGTATTC AACATTTTCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
   TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100
   CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC      150
   AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGCT      200
   GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG      250
20 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG      300
   GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT      350
   AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCTGCCA      400
   ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG      450
   CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT      500
25 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA      550
   TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT      600
   TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC      650
   ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG      700
   GAGCCGGTGA GCGTGGGTCT AGCGGTATCA TTGCAGCACT GGGGCCAGAT      750
30 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC      800
   TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA      850
   AGCATTGGTA A                                     861

```

35

2) INFORMATION FOR SEQ ID NO: 1933

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
 40 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

45

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
 (C) ACCESSION NUMBER: AF190693

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1933

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100

```

CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAGCT GGATCTCAAC 150
AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGGT 200
GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG 250
CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG 300
5 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT 350
AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA 400
ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG 450
CACAACATGG GGGATCATGT AACCCGCCTT GATCGTCGGG AACCGGAGCT 500
GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA 550
10 TGGCAACAAC GTTGCGCAAA CTATTAAGT GCGAACTACT TACTCTAGCT 600
TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC 650
ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG 700
GAGCCGGTGA GCGTGGATCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT 750
GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC 800
15 TATGGATGAA CGAGATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA 850
AGCATTGGTA A 861

20 2) INFORMATION FOR SEQ ID NO: 1934

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
(B) TYPE: Nucleic acid
25 (C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1934

GAACGCCAGC GCGAAATTCA AAAAG

25

35

2) INFORMATION FOR SEQ ID NO: 1935

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
40 (B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1935

AGCTCGGCAT ACTTCGACAG G

21

50

2) INFORMATION FOR SEQ ID NO: 1936

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
5 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1936
10 TACCACCCGC ACGGC 15
- 15 2) INFORMATION FOR SEQ ID NO: 1937
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 bases
 (B) TYPE: Nucleic acid
20 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1937
 CGGAGTCGCC GTCGATG 17
- 30 2) INFORMATION FOR SEQ ID NO: 1938
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 bases
35 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1938
 CCGCGCACCA TTGCTTCGTA CACTGAGGAG TCTCCGCGCG G 41
- 45 2) INFORMATION FOR SEQ ID NO: 1939
- (i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 46 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1939

5

CGACCCGGAT GGTAGTATCG ATAATGATCC GCCAGCGGCC GGGTCG

46

10 2) INFORMATION FOR SEQ ID NO: 1940

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 bases

(B) TYPE: Nucleic acid

15 (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1940

GTATCGTTGG TGACGTAAT

19

25

2) INFORMATION FOR SEQ ID NO: 1941

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases

30 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1941

GCAATGGTCC GTTTAAGT

18

40

2) INFORMATION FOR SEQ ID NO: 1942

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 27 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1942

GACTGGAACA AAGCCTATAA AAAATCA

27

5 2) INFORMATION FOR SEQ ID NO: 1943

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 bases
- (B) TYPE: Nucleic acid
- 10 (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1943

GCTGGTGGAC GGCCAG

16

20

2) INFORMATION FOR SEQ ID NO: 1944

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- 25 (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1944

TTTCGCCGCC ATGCGTTAC

19

35

2) INFORMATION FOR SEQ ID NO: 1945

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 17 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

45 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1945

CGGCGACTAC GCGGTAT

17

50

2) INFORMATION FOR SEQ ID NO: 1946

1013

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 bases
 (B) TYPE: Nucleic acid
5 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1946

CGGCGACTTC GCGGTAT 17
- 15 2) INFORMATION FOR SEQ ID NO: 1947
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 bases
20 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1947

CGGTATACGG CACCATCGT 19
- 30 2) INFORMATION FOR SEQ ID NO: 1948
- (i) SEQUENCE CHARACTERISTICS:
35 (A) LENGTH: 19 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1948

GCGGTATACA ACACCATCG 19
- 45 2) INFORMATION FOR SEQ ID NO: 1949
- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1949

CGGTATACGC CACCATCGT

19

10

2) INFORMATION FOR SEQ ID NO: 1950

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 15 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1950

GGCGACATCG CCTGC

15

25

2) INFORMATION FOR SEQ ID NO: 1951

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 17 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1951

GGCGACAGAG CCTGCTA

17

40

2) INFORMATION FOR SEQ ID NO: 1952

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 19 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1952

CCTGCTATGG AGCGATGGT

19

5

2) INFORMATION FOR SEQ ID NO: 1953

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1953

CGCCTGCTAT AAAGCGATGG T

21

30

2) INFORMATION FOR SEQ ID NO: 1954

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 589 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella pneumoniae* subsp. *pneumoniae*
 (B) STRAIN: ATCC 13883
 (C) ACCESSION NUMBER: AF052258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1954

| | | | | | |
|-------------|------------|------------|------------|------------|-----|
| ACACCGGTCA | ACATTGAGGA | AGAGCTTAAG | AACTCTTATC | TGGATTATGC | 50 |
| GATGTCGGTC | ATTGTTGGCC | GTGCGCTGCC | GGATGTCCGA | GATGGCCTGA | 100 |
| AGCCGGTACA | CCGTCGCGTA | CTTTACGCCA | TGAACGTATT | GGGCAATGAC | 150 |
| TGGAACAAAG | CCTATAAAAA | ATCAGCCCGT | GTCGTTGGTG | ACGTAATCGG | 200 |
| TAAATACCAC | CCGCACGGCG | ACTCCGCGGT | ATACGACACC | ATCGTGCGTA | 250 |
| TGGCGCAGCC | GTTCTCGCTG | CGTTACATGC | TGGTGGACGG | CCAGGGTAAC | 300 |
| TTTGGTTCCA | TCGACGGCGA | CTCCGCCGCG | GCGATGCGTT | ATACCGAAAT | 350 |
| TCGTCTGGCG | AAAATCGCTC | ATGAGCTGAT | GGCCGATCTT | GAAAAAGAGA | 400 |
| CGGTCGATTT | CGTCGACAAC | TATGACGGTA | CGGAGCGTAT | TCCGGACGTC | 450 |
| ATGCCGACCA | AAATTCCTAA | CCTGCTGGTG | AACGGCGCCT | CCGGGATCGC | 500 |
| CGTAGGGGATG | GCCACCAACA | TACCGCCACA | TAACCTGACG | GAAGTGATTA | 550 |
| ACGGCTGTCT | GGCGTATGTT | GACGATGAAG | ACATCAGCA | | 589 |

2) INFORMATION FOR SEQ ID NO: 1955

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1955

CGCGCAACCA TTGCTTCGTA CACTGAGGAG TCTGCGCG

38

2) INFORMATION FOR SEQ ID NO: 1956

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 989 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida inconspicua*
- (B) STRAIN: ATCC 16783

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1956

| | | | | | | |
|----|------------|-------------|------------|------------|------------|-----|
| | CGGTAAACC | ACTTTAACTG | CTGCAATCAC | CAAAGTTTTA | TCTGAAGAAG | 50 |
| | GTGGTGCAGA | TTTCTTAGAT | TATTCATCTA | TTGATAAAGC | ACCAGAAGAG | 100 |
| 35 | AGAGCTAGAG | GTATTACCAT | TTCTACTGCT | CATGTTGAAT | ATGAAACTCC | 150 |
| | AAACAGACAT | TATTCACATG | TTGATTGTCC | AGGTCATCAA | GATTATATTA | 200 |
| | AGAACATGAT | TACTGGTGCA | GCTCAAATGG | ATGGTGCTAT | TATTGTTGTT | 250 |
| | GCAGCTACTG | ATGGTCAAAT | GCCACAAACT | AAGGAACATT | TATTATTAGC | 300 |
| | TAGACAAGTT | GGTGTTC AAC | ATTTAGTTGT | TTTTGTTAAC | AAGTGTGATA | 350 |
| 40 | CCGTTGATGA | TCCAGAAATG | TTRGAATTAG | TTGAAATGGA | AATGAGAGAA | 400 |
| | TTATTGACCG | AATATGGATT | YGATGGTGAT | AACACACCAG | TTATTATGGG | 450 |
| | TTCTGCATTA | ATGGCATTAG | AAGGTAAGAG | ACCAGAAGTT | GGTAAAGAAT | 500 |
| | CTATTGTTAA | GTTAATGGAA | GCTGTTGATA | CTTGATTCC | AACTCCACAA | 550 |
| | AGAGACTTAG | AAAAACCATT | CTTATTACCA | ATTGATGAAG | TTTTCTCTAT | 600 |
| 45 | TTCTGGTAGA | GGTACTGTTG | TTTCTGGTAC | TGTTGATAGA | GGTACATTAA | 650 |
| | AGAAGGGTGA | AGAAGTTGAA | ATTGTTGGTG | GTAAAGAAGG | TGTTATTAAG | 700 |
| | ACTACTGTTA | CCGGTATTGA | AATGTATCAC | AAGGAATTAG | ATCAAGCACA | 750 |
| | AGCTGGTGAT | ACTCCAGGTA | TTTTGTTAAG | AGGTGTTAAG | AGAGATCAAA | 800 |
| | TTGCAAGAGG | TCAAATTCTT | GCAAAGCCAG | GWTCTGTAA | GGCATACAAG | 850 |
| 50 | AAGTTCTTAT | CATCATTATA | CATTTTAAAC | AAGGAAGAAG | GTGGTAGACA | 900 |
| | TACTCCATTT | TCTGAAAATT | ACAGACCTCA | AATGTACATT | AGAACTTCCA | 950 |
| | ATGTTAATGT | TACTTTGAAG | TTCCAGAAA | CTGAAGAAG | | 989 |

2) INFORMATION FOR SEQ ID NO: 1957

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 991 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

10

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida utilis*
 (B) STRAIN: ATCC 22023

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1957

| | | | | | | |
|----|------------|------------|------------|-------------|------------|-----|
| | GGTAAGACCA | CCCTTACTGC | CGCCATCACC | AAGTGCCTTG | CTGAGAAGGG | 50 |
| | AGGTGCCTCG | TTCTTGGA | CTTACAGTGC | CGACAAGGCG | CCAGAGGAGA | 100 |
| | GAGCAAGAGG | TATCACCATC | TCCACTGCGC | ACGTTGAGTA | TGAAACTGCC | 150 |
| 20 | AACAGACACT | ACTCGCACGT | TGACTGTCCA | GGTCACGCTG | ATTACATCAA | 200 |
| | GAACATGATT | ACCGGTGCTG | CGCAGATGGA | CGGTGCTATC | ATTGTCGTTG | 250 |
| | CAGCCACTGA | CGGTCAGATG | CCACAGACCA | GAGAACACTT | GTTGCTTGCC | 300 |
| | AGACAAGTTG | GTGTCCAGCA | CATTGTTGTC | TTCGTCAACA | AGGTTGACAC | 350 |
| | CATCGACGAC | CCTGAGATGC | TTGAGCTTGT | TGAAATGGAG | ATGAGAGAGT | 400 |
| 25 | TGCTTACTTC | GTATGGATTT | GACGGTGATA | ACACCCAGT | TATCATGGGT | 450 |
| | TCTGCTTTGT | GTGCTTTGGA | AGGCCGTGAG | CCAGAGATTG | GTGCTAAGGC | 500 |
| | CATTGACAAG | TTGATGGAGG | CCATTGATGA | GTACATCCCA | ACTCCTCAGA | 550 |
| | GAGACCTGGA | AAAGCCATTC | YTGATGGGTG | TTGAAGACGT | GTTCTCGATC | 600 |
| | TCTGGTAGAG | GTACCGTTGT | CACRGGCCGT | GTTGAGCGTG | GTAACCTGAA | 650 |
| 30 | GAAAGGTGAT | GAAATTGAAC | TTGTTGGCTA | CAACAAGAAC | CCAATCAAGA | 700 |
| | CCACCGTCAC | CGGTATCGAA | ATGTTCAAGA | AGGAGTTGGA | ATCTGCCATG | 750 |
| | GCTGGTGACA | ACTGTGGTAT | CTTGTTGCGT | GGTATCAAGA | GAGATGACGT | 800 |
| | CAAGAGAGGT | ATGGTTGCTG | CTAAGCCAGG | CTCCGTCTCT | GCACACACCA | 850 |
| | AGTTCCTCGC | TTCCTTGAC | ATCCTGACRA | AGGAGGAAGG | TGGTCGTCAC | 900 |
| 35 | AGTGCCTTTG | CTGAGAACTA | CAGACCACAG | ATGTTTCATCA | GAACCGGAGA | 950 |
| | TGTCACCACC | ATCTTGACAT | GGCCAGAGGA | GCACGCTGAC | C | 991 |

40 2) INFORMATION FOR SEQ ID NO: 1958

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 985 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

50 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida zeylanoides*
 (B) STRAIN: ATCC 7351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1958

| | | | | | | |
|----|------------|------------|------------|------------|-------------|-----|
| | CGGTAAGACC | ACTTTGACCG | CCGCCATCAC | CAAGGTGTTG | AGCGCCAAAG | 50 |
| | GTGGTGCTTC | CTTCTTGAC | TACGGGTCCA | TCGACAGAGC | CCCTGAGGAG | 100 |
| 5 | AGAGCCAGAG | GTATTACTAT | CTCGACTGCC | CACGTTGAGT | ACGAGACCGA | 150 |
| | TAAGAGACAC | TACGCCCACG | TTGATTGCCC | TGGTCACGCT | GATTACATCA | 200 |
| | AGAACATGAT | CACTGGTGCC | GCCCAAATGG | ACGGTGCCAT | TATTGTCGTT | 250 |
| | GCTGCTTCTG | ATGGCCAAAT | GCCGCAGACC | AGAGAGCACT | TGTTGCTTGC | 300 |
| | CAGACAGGTT | GGTGTGCAGA | ACTTGGTGTG | GTTTGTTAAC | AAGGTGGACA | 350 |
| 10 | CCATCGACGA | CCCCGAAATG | TTGGAGTTGG | TGGAGATGGA | AATGAGAGAA | 400 |
| | TTGTTGACCC | ACTACGGCTT | TGACGGTGAC | AACACCCCTG | TCATCATGGG | 450 |
| | TTCGGCGTTG | TGTGCCTTGG | AAGACAGGCA | GCCTGAGATT | GGCGAGCAAG | 500 |
| | CCATCATGAA | GTTGTTGGAC | GCTGTCGACG | AGTACATTCC | CACTCCTCAG | 550 |
| | AGAGACTTGG | AGCAACCATT | TTTGATGCCC | GTTGAGGATG | TTTTCTCCAT | 600 |
| 15 | CTCTGGCAGA | GGTACTGTTG | TCACCGGTCG | TGTTGAGAGA | GGCTCATTGA | 650 |
| | AGAAGGGTGA | GGAGATTGAG | ATTGTTGGCG | ACTTCCCCAA | GCCCTTCAAG | 700 |
| | ACTACCGTCA | CCGGCATTGA | GATGTTCAAG | AAGGAGTTGG | ATGCCGCGAT | 750 |
| | GGCGGGCGAC | AACGCCGGGA | TCTTGTTGAG | AGGTGTCAAG | AGAGACGAGG | 800 |
| | TCTCGAGAGG | TATGGTTTTG | GCCAAGCCCG | GTA CTGTAC | TTCGCACACC | 850 |
| 20 | AAGGTGTTGG | CGTCGCTTTA | CATCTTGACC | AAAGAGGAAG | GTGGCCGCCA | 900 |
| | CTCGCCCTTT | GGTGAGAACT | ACAAGCCACA | GTTATTCATC | AGAACCCTCCG | 950 |
| | ATGTCACTGG | TACTTTGAGG | TTCCCCGCCG | GTGAG | | 985 |

25

2) INFORMATION FOR SEQ ID NO: 1959

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 973 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Double
 - (D) TOPOLOGY: Linear

- (ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida catenulata*
- (B) STRAIN: ATCC 10565
- (C) ACCESSION NUMBER:

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1959

| | | | | | | |
|----|------------|------------|------------|-------------|------------|-----|
| | CGGTAAGACC | ACCTTGACTG | CCGCCATCAC | CAAGGTTCTC | TCCGAGAAGG | 50 |
| | GTGGTGCCGA | CTTCTTGAC | TACGGTGCCA | TTGACAGAGC | CCCCGAGGAG | 100 |
| 45 | CGTGCCCGTG | GTATCACCAT | CTCCACTGCC | CACGTTGAGT | ACGAGACTGA | 150 |
| | CAACCGTCAC | TACGCCCACA | TTGACTGTCC | CGGTCACGCT | GATTACATCA | 200 |
| | AGAACATGAT | TACCGGTGCC | GCCCAGATGG | ACGGTGCCAT | TATTGTCCTT | 250 |
| | GCTGCTACTG | ACGGTGCCAT | GCCCCAGACC | CGCGAGCACT | TGCTTCTCGC | 300 |
| | CCGTCAGGTT | GGTATCCAGG | AATTGGTGTG | GTTTGTGAAC | AAGGTTGACA | 350 |
| 50 | CCATCGACGA | CCCCGAGATG | TTGGAGCTCG | TTGAGATGGA | GATCCGCGAG | 400 |
| | TTGTTGTCTG | AGTTCGGTTT | TGACGGTGAC | AACACCCCCG | TCATCATGGG | 450 |
| | TTCCGCTTTG | TGCGCTTTGG | AGGGCAAGCA | GCCCCGAGATT | GGTGAGCAGG | 500 |
| | CTATCACCAA | GTTGATGGCC | GCCGTTGACG | AGCACATCCC | CACCCCCCAG | 550 |

CGTGA CTTGG AGCAGCCTTT CTTGATGCCT GTTGAGGGTG TTTTCTCTAT 600
 CTCTGGCCGT GGTACCGTGG TGA CTTGGTAA GGTGCCCCGT GGTGTCCTCA 650
 AGAAGGGTGA GGAGATTGAG ATTGTTGGCA ACTTTGACAA GCCCTACAAG 700
 GTGA CTTGTA CTGGTATTGA GATGTTCAAG AAGGAGTTGG ACCAGGCCAT 750
 5 GGCTGGTGAC AACGCCGGTA TCTTGTTCG TGGTGTCAAG CGTGACGAGG 800
 TGTCTCGTGG TATGGTTTTG GCCAAGCCCCG GCACTGTTGT CTCGCACAAG 850
 AAGGTTTTGG CTTCGCTTTA CATCTTGACC CAGGAGGAGG GTGGCCGTAA 900
 GACCGGCTTC GGCTCCAAC ACAAGCCCCA GTTGTTCCTG CGCACTACCG 950
 ACGTCACTGG TACCCTCACC TTC 973
 10

2) INFORMATION FOR SEQ ID NO: 1960

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 985 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida krusei*
 25 (B) STRAIN: ATCC 28870

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1960

AAGACTACCT TGA CTTGCTGC AATCACCAAG GTCTTAGCTG ATCAAGGTGG 50
 30 TGCTGATTTT TTAGATTATG CATCTATTGA CAAGGCTCCT GAAGAAAGAG 100
 CAAGAGGTAT TACTATCTCT ACTGCTCACG TTGAGTATGA AACCCCAAAC 150
 AGACATTATT CTCATGTCGA TTGTCCTGGC CATCAAGATT ATATTAAGAA 200
 TATGATTACT GGTGCTGCAC AAATGGATGG TGCTATTATT GTTGTTCGTC 250
 CTACTGATGG TCAAATGCCA CAACTAAGG AACATTTATT ATTAGCAAGA 300
 35 CAAGTTGGTG TTCAACATTT AGTTGTCTTT GTTAATAAAT GTGACACCAT 350
 TGATGACCCA GAAATGTTGG AATTAGTTGA AATGGAAATG AGAGAACTAT 400
 TGTCTGAATA TGGTTTTGAT GGTGATAACA CTCCAGTTAT TATGGGTTCT 450
 GCATTGATGG CTTTAGAAGA CAAGAGACCT GAAGTTGGTA AGGAATCTAT 500
 TTTAAAGTTA ATGGAAGCYG TTGACACATG GATTCCAACC CCAGAGAGAG 550
 10 ATTTAGAAAA ACCATTTTTG TTACCTATTG ATGAAGTTTT CTCAATCTCT 600
 GGTAGAGGTA CTGTCGTTTC TGGTACTGTC GAAAGAGGTA CTTTGAAGAA 650
 GGGTGAAGAA GTTGAAATTG TTGGTGGTAA GGATGGTTCT ATTAAACTA 700
 CTGTCACAGG TATTGAAATG TATCACAAGG AATTAGACCA AGCGCAAGCA 750
 GGTGATACTC CAGGTATTTT ATTAAGAGGT GTCAAGAGAG ACCAAATCAA 800
 15 GAGAGGTCAA ATTTTAGCAA AGCCAGATTC CGTTAAGGCA TACAAGAAGT 850
 TCTTGGCTTC CCTTTATATC TTAACCAAGG AAGAAGGTGG TAGACATACA 900
 CCATTCTCTG AAAACTACAG ACCACAAATG TACATCAGAA CTACCAATGT 950
 TAACGTTACT TTGAAGTTCC CAGACACTGA AGAAG 985

30

2) INFORMATION FOR SEQ ID NO: 1961

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
5 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1961
10 GCTCAAGGCA GATGGCATTC CC 22

15 2) INFORMATION FOR SEQ ID NO: 1962

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 bases
 (B) TYPE: Nucleic acid
20 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1962
 GGACAAGGCG GTTGCGTTTG AT 22

30 2) INFORMATION FOR SEQ ID NO: 1963

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 bases
35 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1963
 CATTCCTGTC TCGCTCGACA GT 22

45 2) INFORMATION FOR SEQ ID NO: 1964

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 19 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1964

5

ATCTGCCTGC CCGTCTTGC

19

10 2) INFORMATION FOR SEQ ID NO: 1965

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 816 bases

(B) TYPE: Nucleic acid

15 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Plasmid pGS05

(C) ACCESSION NUMBER: M36657

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1965

25

| | | | | | |
|---------------|------------|------------|------------|-------------|-----|
| ATGAATAAAT | CGCTCATCAT | TTTCGGCATC | GTCAACATAA | CCTCGGACAG | 50 |
| TTTCTCCGAT | GGAGGCCGGT | ATCTGGCGCC | AGACGCAGCC | ATTGCGCAGG | 100 |
| CGCGTAAGCT | GATGGCCGAG | GGGGCAGATG | TGATCGACCT | CGGTCCGGCA | 150 |
| TCCAGCAATC | CCGACGCCGC | GCCTGTTTCG | TCCGACACAG | AAATCGCGCG | 200 |
| 30 TATCGCGCCG | GTGCTGGACG | CGCTCAAGGC | AGATGGCATT | CCCGTCTCGC | 250 |
| TCGACAGTTA | TCAACCCGCG | ACGCAAGCCT | ATGCCTTGTC | GCGTGGTGTG | 300 |
| GCCTATCTCA | ATGATATTCG | CGGTTTTCCA | GACGCTGCGT | TCTATCCGCA | 350 |
| ATTGGCGAAA | TCATCTGCCA | AACTCGTCGT | TATGCATTCG | GTGCAAGACG | 400 |
| GGCAGGCAGA | TCGGCGCGAG | GCACCCGCTG | GCGACATCAT | GGATCACATT | 450 |
| 35 GCGGCGTTCT | TTGACGCGCG | CATCGCGGCG | CTGACGGGTG | CCGGTATCAA | 500 |
| ACGCAACCGC | CTTGTCCTTG | ATCCCGGCAT | GGGGTTTTTT | CTGGGGGGCTG | 550 |
| CTCCCGAAAC | CTCGCTCTCG | GTGCTGGCGC | GGTTCGATGA | ATTGCGGCTG | 600 |
| CGCTTCGATT | TGCCGGTGCT | TCTGTCTGTT | TCGCGCAAAT | CCTTTCTGCG | 650 |
| CGCGCTCACA | GGCCGTGGTC | CGGGGGATGT | CGGGGCCGCG | ACACTCGCTG | 700 |
| 10 CAGAGCTTGC | CGCCGCCGCA | GGTGGAGCTG | ACTTCATCCG | CACACACGAG | 750 |
| CCGCGCCCCCT | TGCGCGACGG | GCTGGCGGTA | TTGGCGGCGC | TGAAAGAAAC | 800 |
| CGCAAGAATT | CGTTAA | | | | 816 |

15

2) INFORMATION FOR SEQ ID NO: 1966

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 bases

50 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1966

5 CATGCCAGTC TTGCCAACG 19

2) INFORMATION FOR SEQ ID NO: 1967

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

15 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1967

20

CAGCAATAAG TAATCCAGCG ATG 23

25 2) INFORMATION FOR SEQ ID NO: 1968

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

30 (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1968

GGAGAGATTT CACCGCATAG 20

40

2) INFORMATION FOR SEQ ID NO: 1969

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases

45 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1969

AGCCAACCAT CATGCTATTC CA 22

2) INFORMATION FOR SEQ ID NO: 1970

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1206 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

10

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Transposon Tn10
 (C) ACCESSION NUMBER: J01830

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1970

ATGAATAGTT CGACAAAGAT CGCATTGGTA ATTACGTTAC TCGATGCCAT 50
 GGGGATTGGC CTTATCATGC CAGTCTTGCC AACGTTATTA CGTGAATTTA 100
 TTGCTTCGGA AGATATCGCT AACCACCTTG GCGTATTGCT TGCACTTTAT 150
 20 GCGTTAATGC AGGTTATCTT TGCTCCTTGG CTTGGAAAAA TGTCTGACCG 200
 ATTTGGTCGG CGCCAGTGC TGTGTGTGTC ATTAATAGGC GCATCGCTGG 250
 ATTACTTATT GCTGGCTTTT TCAAGTGCGC TTTGGATGCT GTATTTAGGC 300
 CGTTTGCTTT CAGGGATCAC AGGAGCTACT GGGGCTGTCG CGGCATCGGT 350
 CATTGCCGAT ACCACCTCAG CTTCTCAACG CGTGAAGTGG TTCGGTTGGT 400
 25 TAGGGGCAAG TTTTGGGCTT GGTTTAATAG CGGGGCCTAT TATTGGTGGT 450
 TTTGCAGGAG AGATTTCACC GCATAGTCCC TTTTTTATCG CTGCGTTGCT 500
 AAATATTGTC ACTTTCCTTG TGGTTATGTT TTGGTTCCGT GAAACCAAAA 550
 ATACACGTGA TAATACAGAT ACCGAAGTAG GGGTTGAGAC GCAATCGAAT 600
 TCGGTATACA TCACTTTATT TAAAACGATG CCCATTTTGT TGATTATTTA 650
 30 TTTTTCAGCG CAATTGATAG GCCAAATTCC CGCAACGGTG TGGGTGCTAT 700
 TTACCGAAAA TCGTTTGGGA TGGAATAGCA TGATGGTTGG CTTTTCATTA 750
 GCGGGTCTTG GTCTTTTACA CTCAGTATTC CAAGCCTTTG TGGCAGGAAG 800
 AATAGCCACT AAATGGGGCG AAAAAACGGC AGTACTGCTC GAATTTATTG 850
 CAGATAGTAG TGCATTTGCC TTTTATAGCGT TTATATCTGA AGGTTGGTTA 900
 35 GATTTCCCTG TTTTAATTTT ATTGGCTGGT GGTGGGATCG CTTTACCTGC 950
 ATTACAGGGA GTGATGTCTA TCCAAACAAA GAGTCATGAG CAAGGTGCTT 1000
 TACAGGGATT ATTGGTGAGC CTTACCAATG CAACCGGTGT TATTGGCCCA 1050
 TTAAGTTTGA CTGTTATTTA TAATCATTCA CTACCAATTT GGGATGGCTG 1100
 GATTTGGATT ATTGGTTTAG CGTTTACTG TATTATTATC CTGCTATCGA 1150
 40 TGACCTTCAT GTTAACCCCT CAAGCTCAGG GGAGTAAACA GGAGACAAGT 1200
 GCTTAG 1206

15 2) INFORMATION FOR SEQ ID NO: 1971

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 21 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1971

5 CYGACTGYGC CATCCTYATC A

21

2) INFORMATION FOR SEQ ID NO: 1972

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1972

20 MGICAGCTCA TYITTGCWKS C

21

2) INFORMATION FOR SEQ ID NO: 1973

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

30 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1973

35 RACACCRGIY TTGGWITCCT T

21

2) INFORMATION FOR SEQ ID NO: 1974

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
15 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1974

50 ACAAGGGITG GRMSAAGGAG AC

22

2) INFORMATION FOR SEQ ID NO: 1975

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1975

TGRCCRGGGT GGTTRAGGAC G

21

15

2) INFORMATION FOR SEQ ID NO: 1976

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

25 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1976

GATGGAYTCY GTYAAITGGG A

21

30

2) INFORMATION FOR SEQ ID NO: 1977

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1977

15 GATGGAYTCY GTYAARTGGG A

21

2) INFORMATION FOR SEQ ID NO: 1978

50

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
(B) TYPE: Nucleic acid

- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1978

CATCITGYAA TGGYAATCTY AAT

23

10

2) INFORMATION FOR SEQ ID NO: 1979

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1979

CATCYTGYAA TGGYAASCTY AAT

23

25

2) INFORMATION FOR SEQ ID NO: 1980

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1980

40

TCRATGGCIT CIAIRAGRGT YT

22

2) INFORMATION FOR SEQ ID NO: 1981

45

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1981

TGGACACCIS CAAGIGGKCY G

21

5

2) INFORMATION FOR SEQ ID NO: 1982

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1982

TGGACACYIS CAAGIGGKCY G

21

20

2) INFORMATION FOR SEQ ID NO: 1983

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1983

35 CYGAYTGCGC YATICTCATC A

21

2) INFORMATION FOR SEQ ID NO: 1984

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 21 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
45 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1984

50

CYGAYTGYGC YATYCTSATC A

21

2) INFORMATION FOR SEQ ID NO: 1985

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1383 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Cryptococcus neoformans*
 (B) STRAIN: M1-106
 (C) ACCESSION NUMBER: U81804
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1985

| | | | | | | |
|----|------------|------------|------------|------------|------------|------|
| | ATGGGTAAGG | ACAAGCTGCA | CGTCAACGTC | GTTGTTATCG | GTCACGTCGA | 50 |
| | CTCCGGTAAG | TCGACCACCA | CCGGTCACTT | GATCTACAAG | TGCGGTGGTA | 100 |
| 20 | TCGACAAGCG | AACCATTGAG | AAGTTCGAGA | AGGAGGCTCA | AGAGCTCGGA | 150 |
| | AAGTCTTCTT | TCAAGTACGC | TTGGGTTCCT | GACAAGCTTA | AGGCCGAGCG | 200 |
| | AGAGCGAGGT | ATCACCATCG | ACATTGCTCT | TTGGAAGTTC | GAGACCCCTA | 250 |
| | AGTACCAGGT | TACCGTCATT | GACGCCCCCG | GTCACCGAGA | CTTCATCAAG | 300 |
| | AACATGATCA | CCGGTACCTC | CCAGGCTGAC | TGTGCCATCC | TCATCATTCG | 350 |
| 25 | CACCGGTATC | GGTGAGTTCG | AGGCTGGTAT | CTCCAAGGAC | GGTCAGACCC | 400 |
| | GAGAGCACGC | CCTCCTCGCC | TTCACCCTCG | GTGTCAGGCA | GCTCATTGTT | 450 |
| | GCTTGCAACA | AGATGGACAC | CTGCAAGTGG | TCTGAGGACC | GATTCAACGA | 500 |
| | AATCGTCAAG | GAGACCAACG | GTTTCATCAA | GAAGGTTGGT | TACAACCCCA | 550 |
| | AGGCTGTCCC | CTTCGTCCCC | ATCTCTGGTT | GGCACGGTGA | CAACATGTTG | 600 |
| 30 | GAGGAGACCA | CCAACATGCC | CTGGTACAAG | GGATGGACCA | AGGAGACCAA | 650 |
| | GTCCGGTGTT | TCCAAGGGTA | AGACCCTTCT | CGAGGCCATC | GACGCCAGTA | 700 |
| | GGCCCCCTAC | CCGACCCACC | GACAAGCCCC | TCCGTCTCCC | TCTCCAGGAC | 750 |
| | GTCTACAAGA | TCGGTGGTAT | CGGCACAGTC | CCTGTCGGCC | GAGTCGAGAC | 800 |
| | CGGTGTCATC | AAGGCCGGTA | TGGTCGTCAA | GTTTCGCCCC | ACCAACGTCA | 850 |
| 35 | CCACTGAAGT | CAAGTCCGTT | GAGATGCACC | ACGAGCAGAT | CCCCGAGGGT | 900 |
| | CTCCCCGGAG | ACAACGTTGG | TTTCAACGTC | AAGAACGTTT | CCATCAAGGA | 950 |
| | CATCCGACGA | GGTAACGTCT | GTGGTGACTC | CAAGAACGAC | CCCCCTATGG | 1000 |
| | AGGCTGCTTC | TTTCAACGCC | CAGGTTATCG | TCCTTAACCA | CCCTGGTCAG | 1050 |
| | ATCGGTGCCG | GTTACACCCC | CGTTCTCGAC | TGTCACACTG | CCCACATTGC | 1100 |
| 40 | TTGCAAGTTC | TCTGAGTTGA | TCGAGAAGAT | TGACCGACGA | ACCGGTAAGG | 1150 |
| | TCATGGAGGC | CGCCCCCAAG | TTCGTCAAGT | CTGGTGACGC | CGCCATTGTC | 1200 |
| | AAGCTTGTTT | CCCAGAAGCC | TCTCTGTGTT | GAGACCTACG | CCGACTACCC | 1250 |
| | CCCTCTTGGT | CGATTGCGCG | TCCGAGACAT | GCGACAGACC | GTTGCCGTTG | 1300 |
| | GTGTTATTAA | GAGTGTGGAG | AAGTCCGATG | GGAAGAGCGG | CAAGGTTACC | 1350 |
| 45 | AAGGCCGCCG | AGAAGGCTGC | TAAGAAGAAG | TAA | | 1383 |

2) INFORMATION FOR SEQ ID NO: 1986

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1380 bases
 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptococcus neoformans*
(B) STRAIN: B3501
(C) ACCESSION NUMBER: U81803

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1986

| | | | | | | |
|----|------------|------------|------------|------------|-------------|------|
| | ATGGGTAAGG | ACAAGCTGCA | CGTCAACGTC | GTTGTTATCG | GTCACGTCGA | 50 |
| | CTCCGGTAAG | TCGACCACCA | CCGGTCACTT | GATCTACAAG | TGCGGTGGTA | 100 |
| 15 | TCGACAAGCG | AACCATTGAG | AAGTTGCGAG | AGGAGGCTCA | AGAGCTCGGA | 150 |
| | AAGTCTTCTT | TCAAGTACGC | TTGGGTTCCT | GACAAGCTTA | AGGCCGAGCG | 200 |
| | AGAGCGAGGT | ATCACCATCG | ACATTGCTCT | TTGGAAGTTC | GAGACCCCCA | 250 |
| | GGTACCAGGT | CACCGTCATT | GACGCCCCCG | GTCACCGAGA | CTTCATCAAG | 300 |
| | AACATGATCA | CCGGTACCTC | CCAGGCTGAC | TGTGCCATCC | TCATCATTCG | 350 |
| 20 | CACCGGTATC | GGTGAGTTCG | AGGCCGGTAT | CTCCAAGGAC | GGTCAGACCC | 400 |
| | GAGAGCACGC | CCTCCTCGCC | TTCACCCTCG | GTGTCAGGCA | GCTCATTTGT | 450 |
| | GCTTGCAACA | AGATGGACAC | CTGCAAGTGG | TCCGAGGACC | GATTCAACGA | 500 |
| | AATCGTCAAG | GAGACCAACG | GTTTCATCAA | GAAGGTTGGC | TACAACCCCCA | 550 |
| | AGGCTGTCCC | CTTCGTCCCC | ATCTCTGGTT | GGCACGGTGA | CAACATGTTG | 600 |
| 25 | GAGGAGACCA | CCAACATGCC | CTGGTACAAG | GGATGGACCA | AGGAGACCAA | 650 |
| | GTCTGGTGTG | TCCAGGGGTA | AGACCCTTCT | CGAGGCCATC | AGCGCCAGTA | 700 |
| | GGCCCCATAC | CCGACCCACC | GACAAGCCCC | TCCGTCTCCC | TCTCCAGGAC | 750 |
| | GTCTACAAGA | TCGGTGGTAT | CGGCACAGTC | CCTGTCGGCC | GAGTCGAGAC | 800 |
| | CGGTGTCATC | AAGGCCGGTA | TGGTCGTCAA | GTTGCGCCCC | ACCAACGTCA | 850 |
| 30 | CCACTGAAGT | CAAGTCCGTT | GAGATGCACC | ACGAGCAGAT | CCCCGAGGGT | 900 |
| | CTTCCCGGAG | ACAACGTTGG | TTTCAACGTC | AAGAACGTTT | CCATCAAGGA | 950 |
| | CATCCGACGA | GGTAACGTCT | GTGGTGACTC | CAAGAACGAC | CCCCCTATGG | 1000 |
| | AGGCTGCTTC | TTTCAACGCC | CAGGTTATCG | TCCTTAACCA | CCCTGGTCAG | 1050 |
| | ATCGGTGCCG | GTTACACCCC | CGTTCTCGAC | TGTCACACTG | CCCACATTGC | 1100 |
| 35 | CTGCAAGTTT | GCTGAGTTGA | TCGAGAAGAT | TGACCGACGA | ACCGGTAAGG | 1150 |
| | TCATGGAGGC | CGCCCCAAG | TTCGTCAAGT | CTGGTGACGC | CGCCATTGTC | 1200 |
| | AAGCTTGTTG | CCCAGAAGCC | CCTCTGTGTT | GAGACCTACG | CCGACTACCC | 1250 |
| | CCCTCTTGGT | CGATTGCGCG | TCCGAGACAT | GCGACAGACC | GTTGCCGTTG | 1300 |
| | GTGTTATCAA | GAGCGTGGAC | AAGACCGAGA | AGGGTGGCAA | GGTCACCAAG | 1350 |
| 40 | GCTGCTGAGA | AGGCTGCCAA | GAAGAAGTAA | | | 1380 |

2) INFORMATION FOR SEQ ID NO: 1987

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1377 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(C) ACCESSION NUMBER: X01638

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1987

```

ATGGGTAAAG AGAAGTCTCA CATTACGTT GTCGTTATCG GTCATGTCGA      50
TTCTGGTAAG TCTACCACTA CCGGTCATTT GATTTACAAG TGTGGTGGTA      100
TTGACAAGAG AACCATCGAA AAGTTCGAAA AGGAAGCCGC TGAATTAGGT      150
10 AAGGGTTCTT TCAAGTACGC TTGGGTTTTG GACAAGTTAA AGGCTGAAAG      200
AGAAAGAGGT ATCACTATCG ATATTGCTTT GTGGAAGTTC GAAACTCCAA      250
AGTACCAAGT TACCGTTATT GATGCTCCAG GTCACAGAGA TTTCATCAAG      300
AACATGATTA CTGGTACTTC TCAAGCTGAC TGTGCTATCT TGATTATTGC      350
TGGTGGTGTC GGTGAATTCG AAGCCGGTAT CTCTAAGGAT GGTCAAACCA      400
15 GAGAACACGC TTTGTTGGCT TTCACCTTGG GTGTTAGACA ATTGATTGTT      450
GCTGTCAACA AGATGGACTC CGTCAAATGG GACGAATCCA GATTCCAAGA      500
AATTGTCAAG GAAACCTCCA ACTTTATCAA GAAGGTTGGT TACAACCCAA      550
AGACTGTTCC ATTCGTCCCA ATCTCTGGTT GGAACGGTGA CAACATGATT      600
GAAGCTACCA CCAACGCTCC ATGGTACAAG GGTGGGAAA AGGAAACCAA      650
20 GGCCGGTGTC GTCAAGGGTA AGACTTTGTT GGAAGCCATT GACGCCATTG      700
AACAACCATC TAGACCAACT GACAAGCCAT TGAGATTGCC ATTGCAAGAT      750
GTTTACAAGA TCGGTGGTAT TGGTACTGTG CCAGTCGGTA GAGTTGAAAC      800
CGGTGTCATC AAGCCAGGTA TGGTTGTTAC TTTCGCCCCA GCTGGTGTTA      850
CCACTGAAGT CAAGTCCGTT GAAATGCATC ACGAACAATT GGAACAAGGT      900
25 GTTCCAGGTG ACAACGTTGG TTTCAACGTC AAGAACGTTT CCGTTAAGGA      950
AATCAGAAGA GGTAACGTCT GTGGTGACGC TAAGAACGAT CCACCAAAGG     1000
GTTGCGCTTC TTTCAACGCT ACCGTCATTG TTTTGAACCA TCCAGGTCAA     1050
ATCTCTGCTG GTTACTCTCC AGTTTTGGAT TGTCACACTG CTCACATTGC     1100
TTGTAGATTC GACGAATTGT TGGAAAAGAA CGACAGAAGA TCTGGTAAGA     1150
30 AGTTGGAAGA CCATCCAAAG TTCTTGAAGT CCGGTGACGC TGCTTTGGTC     1200
AAGTTCGTTC CATCTAAGCC AATGTGTGTT GAAGCTTTCA GTGAATACCC     1250
ACCATTAGGT AGATTGCTG TCAGAGACAT GAGACAAACT GTCGCTGTGCG     1300
GTGTTATCAA GTCTGTTGAC AAGACTGAAA AGGCCGCTAA GGTTACCAAG     1350
GCTGCTCAAA AGGCTGCTAA GAAATAA                                     1377
35

```

2) INFORMATION FOR SEQ ID NO: 1988

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1377 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

50 (C) ACCESSION NUMBER: M10992

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1988

| | | | | | | |
|----|------------|------------|------------|------------|------------|------|
| | ATGGGTAAG | AGAAGTCTCA | CATTAACGTT | GTCGTTATCG | GTCATGTCGA | 50 |
| | TTCTGGTAAG | TCTACCACTA | CCGGTCATTT | GATTTACAAG | TGTGGTGGTA | 100 |
| | TTGACAAGAG | AACCATCGAA | AAGTTGCGAA | AGGAAGCCGC | TGAATTAGGT | 150 |
| | AAGGGTTCTT | TCAAGTACGC | TTGGGTTTTG | GACAAGTTAA | AGGCTGAAAG | 200 |
| 5 | AGAAAGAGGT | ATCACTATCG | ATATTGCTTT | GTGGAAGTTC | GAAACTCCAA | 250 |
| | AGTACCAAGT | TACCGTTATT | GATGCTCCAG | GTCACAGAGA | TTTCATCAAG | 300 |
| | AACATGATTA | CTGGTACTTC | TCAAGCTGAC | TGTGCTATCT | TGATTATTGC | 350 |
| | TGGTGGTGTC | GGTGAATTCG | AAGCCGGTAT | CTCTAAGGAT | GGTCAAACCA | 400 |
| | GAGAACACGC | TTTGTTGGCT | TTCACCTTGG | GTGTTAGACA | ATTGATTGTT | 450 |
| 10 | GCTGTCAACA | AGATGGACTC | CGTCAAATGG | GACGAATCCA | GATTCCAAGA | 500 |
| | AATTGTCAAG | GAAACCTCCA | ACTTTATCAA | GAAGGTTGGT | TACAACCCAA | 550 |
| | AGACTGTTCC | ATTCGTCCCA | ATCTCTGGTT | GGAACGGTGA | CAACATGATT | 600 |
| | GAAGCTACCA | CCAACGCTCC | ATGGTACAAG | GGTTGGGAAA | AGGAAACCAA | 650 |
| | GGCCGGTGTC | GTCAAGGGTA | AGACTTTGTT | GGAAGCCATT | GACGCCATTG | 700 |
| 15 | AACAACCATC | TAGACCAACT | GACAAGCCAT | TGAGATTGCC | ATTGCAAGAT | 750 |
| | GTTTACAAGA | TTGGTGGTAT | TGGTACTGTG | CCAGTCGGTA | GAGTTGAAAC | 800 |
| | CGGTGTCATC | AAGCCAGGTA | TGGTTGTTAC | TTTTGCCCCA | GCTGGTGTTA | 850 |
| | CCACTGAAGT | CAAGTCCGTT | GAAATGCATC | ACGAACAATT | GGAACAAGGT | 900 |
| | GTTCCAGGTG | ACAACGTTGG | TTTCAACGTC | AAGAACGTTT | CCGTTAAGGA | 950 |
| 20 | AATCAGAAGA | GGTAACGTCT | GTGGTGACGC | TAAGAACGAT | CCACCAAAGG | 1000 |
| | GTTGCGCTTC | TTTCAACGCT | ACCGTCATTG | TTTTGAACCA | TCCAGGTCAA | 1050 |
| | ATCTCTGCTG | GTTACTCTCC | AGTTTTGGAT | TGTCACACTG | CTCACATTGC | 1100 |
| | TTGTAGATTC | GACGAATTGT | TGGAAAAGAA | CGACAGAAGA | TCTGGTAAGA | 1150 |
| | AGTTGGAAGA | CCATCCAAAG | TTCTTGAAGT | CCGGTGACGC | TGCTTTGGTC | 1200 |
| 25 | AAGTTCGTTC | CATCTAAGCC | AATGTGTGTT | GAAGCTTTCA | GTGAATACCC | 1250 |
| | ACCATTAGGT | AGATTGCTG | TCAGAGACAT | GAGACAAACT | GTCGCTGTCG | 1300 |
| | GTGTTATCAA | GTCTGTTGAC | AAGACTGAAA | AGGCCGCTAA | GGTTACCAAG | 1350 |
| | GCTGCTCAAA | AGGCTGCTAA | GAAATAA | | | 1377 |

30

2) INFORMATION FOR SEQ ID NO: 1989

- (i) SEQUENCE CHARACTERISTICS:
- 35 (A) LENGTH: 1377 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- 40 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Eremothecium gossypii*
- (B) STRAIN: ATCC 10895
- 45 (C) ACCESSION NUMBER: X73978

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1989

| | | | | | | |
|----|------------|------------|------------|------------|------------|-----|
| | ATGGGTAAGG | AAAAGACTCA | CGTTAACGTT | GTCGTCATCG | GTCACGTCGA | 50 |
| 50 | CTCTGGTAAG | TCTACTACCA | CCGGTCACTT | GATCTACAAG | TGTGGTGGTA | 100 |
| | TTGACAAGAG | AACCATCGAG | AAGTTGCGAG | AGGAGGCTGC | CGAGTTGGGT | 150 |
| | AAGGGTTCTT | TCAAGTACGC | CTGGGTTTTG | GACAAATTGA | AGGCTGAGAG | 200 |
| | AGAGAGAGGT | ATCACCATCG | ACATTGCGTT | GTGGAAGTTC | GAGACTCCAA | 250 |

| | | | | | | |
|----|------------|------------|------------|------------|------------|------|
| | AGTACCACGT | CACTGTCATT | GACGCCCCAG | GCCACAGAGA | CTTCATCAAG | 300 |
| | AACATGATTA | CCGGTACTTC | TCAAGCTGAC | TGTGCCATCT | TGATCATTGC | 350 |
| | TGGTGGTGTC | GGTGAGTTCG | AGGCTGGTAT | CTCCAAGGAC | GGTCAGACCA | 400 |
| | GAGAGCACGC | TTTGTTGGCT | TACACCTTGG | GTGTCAAGCA | GTTGATCGTT | 450 |
| 5 | GCCATCAACA | AGATGGACTC | CGTCAAGTGG | GACGAGTCCA | GATACCAGGA | 500 |
| | GATTGTCAAG | GAGACCTCCA | ACTTCATCAA | GAAGGTCGGT | TACAACCCTA | 550 |
| | AGACTGTTCC | ATTCGTTCCA | ATCTCCGGCT | GGAACGGTGA | CAACATGATT | 600 |
| | GAGGCCACCA | CCAACGCCCC | ATGGTACAAG | GGCTGGGAGA | AGGAGACCAA | 650 |
| | GGCTGGTGCC | GTCAAGGGTA | AGACCTTGTT | GGAGGCCATT | GACGCCATTG | 700 |
| 10 | AGCCACCTGT | CAGACCAACT | GACAAGGCAT | TGAGATTGCC | ATTGCAGGAT | 750 |
| | GTCTACAAGA | TCGGTGGTAT | TGGTACGGTT | CCAGTCGGCA | GAGTCGAGAC | 800 |
| | CGGTGTCATC | AAGCCAGGTA | TGGTTGTTAC | CTTCGCCCCA | TCCGGTGTCA | 850 |
| | CCACTGAAGT | CAAGTCCGTC | GAGATGCACC | ACGAGCAATT | GGAGGAGGGT | 900 |
| | GTCCCAGGTG | ACAACGTTGG | TTTCAACGTC | AAGAACGTCT | CCGTCAAGGA | 950 |
| 15 | GATCAGAAGA | GGTAACGTTT | GCGGTGACTC | CAAGAACGAC | CCACCAAAGG | 1000 |
| | CTGCTGAGTC | CTTCAACGCT | ACCGTCATTG | TCTTGAACCA | CCCAGGTCAA | 1050 |
| | ATCTCTGCCG | GTTACTCTCC | AGTCTTGGAC | TGTCACACTG | CCCACATTGC | 1100 |
| | TTGTAAGTTC | GACGAGTTGT | TGGAGAAGAA | CGACAGAAGA | ACCGGTAAGA | 1150 |
| | AGTTGGAAGA | CTCTCCAAAG | TTCCTAAAGG | CCGGTGACGC | TGCCATGGTC | 1200 |
| 20 | AAGTTTGTCC | CATCCAAGCC | AATGTGTGTT | GAGGCTTTCA | CCGACTACCC | 1250 |
| | ACCATTGGGT | AGATTGCTG | TCAGAGACAT | GAGACAGACC | GTTGCTGTCG | 1300 |
| | GTGTCATCAA | GTCTGTTGTC | AAGTCCGACA | AGGCTGGTAA | GGTCACCAAG | 1350 |
| | GCCGCCAAA | AGGCTGGTAA | GAAATAG | | | 1377 |

25

2) INFORMATION FOR SEQ ID NO: 1990

- (i) SEQUENCE CHARACTERISTICS:
- 30 (A) LENGTH: 1377 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- 35 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Eremothecium gossypii*
- (C) ACCESSION NUMBER: A29820
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1990

| | | | | | | |
|----|------------|------------|------------|------------|------------|-----|
| | ATGGGTAAGG | AAAAGACTCA | CGTTAACGTT | GTCGTCATCG | GTCACGTCGA | 50 |
| | CTCTGGTAAG | TCTACTACCA | CCGGTCACTT | GATCTACAAG | TGTGGTGGTA | 100 |
| 15 | TTGACAAGAG | AACCATCGAG | AAGTTCGAGA | AGGAGGCTGC | CGAGTTGGGT | 150 |
| | AAGGGTTCTT | TCAAGTACGC | CTGGGTTTTG | GACAAATTGA | AGGCTGAGAG | 200 |
| | AGAGAGAGGT | ATCACCATCG | ACATTGCGTT | GTGGAAGTTC | GAGACTCCAA | 250 |
| | AGTACCACGT | CACTGTCATT | GACCCCCCAG | GCCACAGAGA | CTTCATCAAG | 300 |
| | AACATGATTA | CCGGTACTTC | TCAAGCTGAC | TGTGCCATCT | TGATCATTGC | 350 |
| 50 | TGGTGGTGTC | GGTGAGTTCG | AGGCTGGTAT | CTCCAAGGAC | GGTCAGACCA | 400 |
| | GAGAGCACGC | TTTGTTGGCT | TACACCTTGG | GTGTCAAGCA | GTTGATCGTT | 450 |
| | GCCATCAACA | AGATGGACTC | CGTCAAGTGG | GACGAGTCCA | GATACCAGGA | 500 |
| | GATTGTCAAG | GAGACCTCCA | ACTTCATCAA | GAAGGTCGGT | TACAACCCTA | 550 |

| | | | | | | |
|----|------------|-------------|------------|------------|-------------|------|
| | AGACTGTTCC | ATTTCGTTCCA | ATCTCCGGCT | GGAACGGTGA | CAACATGATT | 600 |
| | GAGGCCACCA | CCAACGCCCC | ATGGTACAAG | GGCTGGGAGA | AGGAGACCAA | 650 |
| | GGCTGGTGCC | GTCAAGGGTA | AGACCTTGTT | GGAGGCCATT | GACGCCATTG | 700 |
| | AGCCACCTGT | CAGACCAACT | GACAAGGCAT | TGAGATTGCC | ATTGCAGGAT | 750 |
| 5 | GTCTACAAGA | TCGGTGGTAT | TGGTACGGTT | CCAGTCGGCA | GAGTCGAGAC | 800 |
| | CGGTGTCATC | AAGCCAGGTA | TGGTTGTTAC | CTTCGCCCCA | TCCGGTGTCA | 850 |
| | CCACTGAAGT | CAAGTCCGTC | GAGATGCACC | ACGAGCAATT | GGAGGAGGGT | 900 |
| | GTCCCAGGTG | ACAACGTTGG | TTTCAACGTC | AAGAACGTCT | CCGTCAAGGA | 950 |
| | GATCAGAAGA | GGTAACGTTT | GCGGTGACTC | CAAGAACGAC | CCACCAAAGG | 1000 |
| 10 | CTGCTGAGTC | CTTCAACGCT | ACCGTCATTG | TCTTGAACCA | CCCAGGTCAA | 1050 |
| | ATCTCTGCCG | GTTACTCTCC | AGTCTTGAC | TGTCACACTG | CCCACATTGC | 1100 |
| | TTGTAAGTTC | GACGAGTTGT | TGGAGAAGAA | CGACAGAAGA | ACCGGTAAGA | 1150 |
| | AGTTGGAAGA | CTCTCCAAAG | TTCCTAAAGG | CCGGTGACGC | TGCCATGGTC | 1200 |
| | AAGTTTGTCC | CATCCAAGCC | AATGTGTGTT | GAGGCTTTCA | CCGACTACCC | 1250 |
| 15 | ACCATTGGGT | AGATTTCGCTG | TCAGAGACAT | GAGACAGACC | GTTGCTGTCTG | 1300 |
| | GTGTCATCAA | GTCTGTTGTC | AAGTCCGACA | AGGCTGGTAA | GGTCACCAAG | 1350 |
| | GCCGCCCAA | AGGCTGGTAA | GAAATAG | | | 1377 |

20

2) INFORMATION FOR SEQ ID NO: 1991

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1646 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

30

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus oryzae*
 (B) STRAIN: KBN616
 (C) ACCESSION NUMBER: AB007770

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1991

| | | | | | | |
|----|-------------|------------|------------|-------------|-------------|-----|
| | TGGGGTAAGT | TTATCAACCC | GTCGAGTTGT | GTTGCATCTC | AGATCATGGC | 50 |
| | TGACAAGTAC | TTTCCTCCTT | ACAGTAAGGA | AGACAAGCAG | CACATCAACA | 100 |
| 40 | TCGTGCTTAT | CGGCCACGTC | GATTCCGGCA | AGTCCACCAC | CACTGGTCAC | 150 |
| | TTGATCTACA | AGTGTGGTGG | TATCGACCAG | CGTACCATCG | AGAAGTTCGA | 200 |
| | GAAGGAAGCC | GCTGAGCTCG | GTAAGGGTTC | CTTCAAGTAC | GCCTGGGTTC | 250 |
| | TTGACAAGCT | CAAGTCCGAG | CGTGAGCGTG | GTATCACCAT | CGATATCGCC | 300 |
| | CTCTGGAAGT | TCCAGACCTC | CAAGTATGAG | GTCACCGTCA | TTGGTAAGCA | 350 |
| 45 | TTTGAGTTCC | AACCTACGTT | GCCCAACATT | TACAGTCATC | TAACAAAGTT | 400 |
| | CAATAGATGC | CCCCGGTCAC | CGTGACTTCA | TCAAGAACAT | GATCACTGGT | 450 |
| | ACTTCCAGG | CTGACTGCGC | TATCCTCATC | ATTGCCCTCCG | GTAAGTGGTGA | 500 |
| | ATTGAGAGGCT | GGTATCTCCA | AGGATGGTCA | GACCCGTGAG | CACGCTCTGC | 550 |
| | TCGCTTTTAC | CCTCGGTGTC | CGTCAGCTCA | TCGTTGCCCT | CAACAAGATG | 600 |
| 50 | GACACCTGCA | AGTGGTCTCA | GGATCGTTAC | AACGAAATCG | TTAAGGAGAC | 650 |
| | TTCCAACCTC | ATCAAGAAGG | TCGGATACAA | CCCCAAGAGC | GTTCTTTTCG | 700 |
| | TCCCCATCTC | CGGTTTCAAC | GGTGACAACA | TGATTGAGGC | CTCCACCAAC | 750 |
| | TGCCCCCTGGT | ACAAGGGCTG | GGAGAAGGAG | ACCAAGGCTG | GCAAGTCCAC | 800 |

| | | | | | | |
|----|------------|------------|------------|------------|-------------|------|
| | CGGTAAGACC | CTTCTCGAGG | CCATCGATGC | CATCGAGCCC | CCCGTCCGTC | 850 |
| | CCACCGACAA | GCCTCTCCGT | CTTCCCCTCC | AGGATGTCTA | CAAGATCTCT | 900 |
| | GGTATCGGTA | CTGTGCCCCG | CGGTCGTGTC | GAGACTGGTG | TCATCAAGCC | 950 |
| | TGGTATGGTC | GTTACTTTTC | CTCCTGCCAA | CGTGACCACT | GAAGTCAAGT | 1000 |
| 5 | CCGTTGAAAT | GCACCACCAG | CAGCTCCAGG | CCGGTAACCC | CGGTGACAAC | 1050 |
| | GTTGGTTTCA | ACGTCAAGAA | CGTCTCCGTC | AAGGAAGTCC | GCCGTGGTAA | 1100 |
| | CGTTGCCGGT | GACTCCAAGA | ACGACCCCCC | TGCTGGCTGC | GATTCCTTCA | 1150 |
| | ACGCCCAGGT | CATCGTCCTT | AACCACCCCG | GTCAGGTCGG | CAACGGTTAC | 1200 |
| | GCTCCCGTCC | TGGACTGCCA | CACCGCTCAC | ATTGCTTGCA | AGTTCGCTGA | 1250 |
| 10 | GCTCCTTGAG | AAGATTGACC | GCCGTACCGG | TAAATCTGTT | GAGGACAAGC | 1300 |
| | CCAAGTTCAT | CAAGTCTGGT | GATGCTGCCA | TCGTCAAGAT | GATTCCTTCC | 1350 |
| | AAGCCCATGT | GTGTGGAGTC | TTTCACTGAC | TTCCCCCCTC | TTGGTCGTTT | 1400 |
| | CGCTGTCCGT | GACGTAAGTT | TTTCCCTCTT | GACTATCTTC | ACAATTTTTTC | 1450 |
| | ACATATTTTC | ACGCCTCGTC | CCACTCTTTT | TCCTCCCTTC | CTCTTTGGTT | 1500 |
| 15 | CCCCTTTTTG | CCTGCAAGTT | CTCTATAGCT | AACATGATGT | CTAGATGCGT | 1550 |
| | CAAACGTGTT | CCGTCGGAGT | TATCAAGTCG | GTTGAGAAGA | ACACTGGCGG | 1600 |
| | TTCTGGCAAG | GTCACCAAGG | CCGCCCAGAA | GGCTGGCAAG | AAATAA | 1646 |

20

2) INFORMATION FOR SEQ ID NO: 1992

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1380 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

30

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aureobasidium pullulans*
- (B) STRAIN: R106
- (C) ACCESSION NUMBER: U19723

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1992

| | | | | | | |
|----|------------|------------|------------|------------|-------------|-----|
| | ATGGGTAAGG | AAAAGTCCCA | CATCAACGTC | GTCGTTATCG | GCCACGTCGA | 50 |
| | CTCCGGTAAG | TCGACCACCA | CCGGTCACTT | GATCTACAAG | TGCGGTGGTA | 100 |
| 40 | TCGACAAGCG | TACCATCGAG | AAGTTCGAGA | AGGAAGCCGC | CGAACTCGGC | 150 |
| | AAGGGTTTCT | TCAAGTACGC | CTGGGTCCTC | GACAAGCTGA | AGTCTGAGCG | 200 |
| | TGAGCGTGGT | ATCACTATCG | ATATCGCTCT | GTGGAAGTTC | GAGACCCCCA | 250 |
| | AGTACATGGT | CACCGTCATC | GATGCCCCCG | GTCACCGTGA | TTTCATCAAG | 300 |
| | AACATGATCA | CTGGTACCTC | CCAGGCTGAC | TGCGCCATTC | TCATCATTGC | 350 |
| 45 | CGCCGGTACT | GGTGAGTTCG | AGGCTGGTAT | CTCCAAGGAT | GGCCAGACTC | 400 |
| | GTGAGCACGC | CCTTCTCGCC | TACACCCTTG | GTGTCAAGCA | GCTCATCGTC | 450 |
| | GCTATCAACA | AGATGGACAC | CACCAAGTGG | TCTGAGGCC | GTTACCAGGA | 500 |
| | GATCATCAAG | GAGACCTCCG | GTTTCATCAA | GAAGGTCGGC | TACAACCCCCA | 550 |
| | AGCACGTTCC | CTTTGTCCCC | ATCTCGGGTT | TCAACGGTGA | CAACATGATT | 600 |
| 50 | GAGGTTTCTT | CCAACTGCCC | CTGGTACAAG | GGTTGGGAGA | AGGAGACCAA | 650 |
| | GGCCAAGGCC | ACTGGTAAGA | CTCTCCTCGA | GGCCATTGAC | GCCATCGACC | 700 |
| | CTCCTTCGCG | CCCCACCGAC | AAGCCCCCTC | GTCTTCCCCT | CCAGGATGTC | 750 |
| | TACAAGATCG | GTGGTATTGG | CACGGTGCCC | GTCGGCCGTG | TCGAGACCGG | 800 |

TACCATCAAG GGTGGTATGG TCGTCACCTT CGCCCCCGCT GGTGTCACCA 850
 CTGAGGTCAA GTCCGTCGAG ATGCACCACG AGCAGCTCTC CGAGGGTCTC 900
 CCCGGTGACA ACGTCGGGCTT CAACGTCAAG AACGTCTCCG TCAAGGAGAT 950
 CCGTCGTGGT AACGTTGCCG GTGACTCCAA GAACGACCCC CCCAAGGGTT 1000
 5 GCGACTCCTT CAACGCCCAG GTCATCGTCC TCAACCACCC CGGTCAGGTC 1050
 GGTGCTGGTT ACGCACCCGT CCTCGATTGC CACACTGCCC ACATCGCCTG 1100
 CAAGTTCTCC GAGCTTGTTG AGAAGATTGA CCGCCGTACC GGCAAGTCCG 1150
 TTGAGGCCCG CCCCAAGTTC ATCAAGTCTG GTGACGCCGC CATCGTCAAG 1200
 ATGGTTCCCT CCAAGCCTAT GTGTGTTGAG GCCTTCACTG ACTACCCTCC 1250
 10 TCTCGGTCGT TTCGCCGTCC GTGACATGAG ACAGACCGTC GCTGTCGGTG 1300
 TCATCAAGTC CGTCGCCAAG TCCGACAAGC AGGGTGCCGG TAAGGTTACC 1350
 AAGGCCGCTG TCAAGGCTGG CAAGAAGTAA 1380

15

2) INFORMATION FOR SEQ ID NO: 1993

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1383 bases
 20 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Histoplasma capsulatum*
 (B) STRAIN: 186AS
 (C) ACCESSION NUMBER: U14100

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1993

ATGGGTAAGG AAGACAAGAC TCACATTAAC CTCGTCGTCA TCGGCCACGT 50
 CGATTCCGGC AAATCTACCA CCACTGGTCA TTTGATCTAC AAATGCGGTG 100
 35 GTATTGACAG CCGTACCATT GAGAAGTTCG AAAAGGAAGC CGAAGAGTTG 150
 GGCAAGAAAT CCTTCAAATA TGCGTGGGTC CTTGACAAAC TGAAGTCTGA 200
 GCGTGAGCGT GGTATCACCA TCGATATTGC CCTCTGGAAA TTCGAGACTC 250
 CGAAGTACAG TGTCACGTGC ATTGATGCTC CCGGCCATCG TGACTTCATC 300
 AAGAACATGA TCACTGGTAC CTCCCAGGCT GACTGCGCTA TCCTCATCAT 350
 40 TGCTGCCGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA 400
 CTCGTGAGCA CGCTCTGCTT GCTTTCACCC TTGGTGTGAG GCAACTCATC 450
 GTTGCCATCA ACAAGATGGA CACCACCAAG TGGTCCGAGT CCCGTTTCAA 500
 CGAAATCATC AAGGAGGTTT CCAACTTCAT CAAGAAGGTC GGATATAACC 550
 CCAAGGCTGT TCCCTTCGTG CCAATCTCTG GTTTCGAGGG TGACAACATG 600
 45 ATTGAACCCT CCCCCAAGT CACATGGTAC AAGGGCTGGA ACAAGGAGAC 650
 TGCCTCTGGC AAGTCTTCTG GTAAAACCCT TCTCGATGCC ATTGACGCCA 700
 TTGAACCCCC AACCCGTCCT ACCGATAAGC CCCTCCGTCT TCCCCTCCAG 750
 GATGTGTACA AAATCTCTGG TATTGGCACT GTTCCCGTCG GACGTGTTGA 800
 GACTGGTGTC ATCAAGCCCG GTATGGTCGT GACTTTCGCT CCCTCCAACG 850
 50 TCACCACTGA AGTCAAGTCC GTCGAAATGC ACCACCAACA ACTCCAGGCT 900
 GGTTACCCTG GCGACAACGT CGGCTTCAAC GTCAAGAACG TTTCAGTCAA 950
 GGAAGTCCGC CGTGGCAACG TTGCTGGCGA CTCCAAAAT GATCCCCCTA 1000
 AGGGCTGCGA ATCCTTCAAT GCCCAGGTCA TCGTCCTTAA CCACCCCGGC 1050

CAGGTTGGCG CTGGTTATGC CCCAGTCCTC GACTGCCACA CTGCCCACAT 1100
 TGCTTGCAAG TTCTCTGAGC TTATTGAGAA GATCGACCGC CGTACCGGAA 1150
 AGTCTGTTGA GAACAACCCC AAGTTCATCA AGTCTGGTGA TGCTGCTATC 1200
 GTCAAGATGG TTCCCTCCAA GCCCATGTGC GTGGAGCCCT TCACTGACTA 1250
 5 TCCCCCTCTT GGACGTTTCG CTGTCCGTGA CATGAGACAA ACCGTCGCTG 1300
 TCGGTGTCAT CAAGTCCGTC ATCAAGTCTG ACAAGACTGC TGGCAAGGTC 1350
 ACCAAGGCCG CGCAGAAGGC CACCAAGAAA TAA 1383

10

2) INFORMATION FOR SEQ ID NO: 1994

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1383 bases
- 15 (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- 20 (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Neurospora crassa*
- (C) ACCESSION NUMBER: D45837

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1994

ATGGGCAAGG AGGACAAGAC TCACATCAAC GTCGTCGTTA TCGGCCACGT 50
 CGATTCCGGC AAGTCTACCA CTACCGGTCA CTTGATCTAC AAGTGCGGTG 100
 GTATCGACAA GCGTACCATC GAGAAGTTCG AGAAGGAAGC CGCTGAGCTC 150
 30 GGTAAGGGTT CCTTCAAGTA TGCCTGGGTT CTTGACAAGC TCAAGGCCGA 200
 GCGTGAGCGT GGTATCACCA TCGATATCGC CCTCTGGAAG TTCGAGACTC 250
 CCAAGTACTA CGTCACCGTC ATCGATGCCC CCGGTCATCG TGATTTTCATC 300
 AAGAACATGA TCACTGGTAC CTCCCAGGCT GATTGCGCTA TCCTCATCAT 350
 TGCCGCCGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA 400
 35 CCCGTGAGCA CGCCCTGCTC GCCTACACCC TCGGTGTCAA GCAGCTCATT 450
 GTTGCCATCA ACAAGATGGA CACCACCCAG TGGTCCCAGA CTCGTTTCGA 500
 GGAGATCATC AAGGAGACCA AGAACTTCAT CAAGAAGGTT GGCTACAACC 550
 CCGCTGGTGT CGCTTTCGTC CCCATCTCCG GCTTCAACGG CGACAACATG 600
 CTTGAGCCCT CCACCAACTG CCCCTGGTAC AAGGGTTGGG AGAAGGAGAC 650
 40 CAAGGCCGGC AAGGCCACTG GCAAGACCCT CCTCGAGGCC ATCGACGCCA 700
 TTGAGCCCCC CAAGCGTCCT ACCGACAAGC CCCTCCGTCT TCCCCTCCAG 750
 GATGTCTACA AGATCGGTGG TATCGGCACA GTGCCCGTCG GCCGTATCGA 800
 GACTGGTGTC CTCAAGCCCG GTATGGTCGT TACCTTCGCT CCTTCCAACG 850
 TCACCACTGA AGTCAAGTCC GTCGAGATGC ACCACGAGCA GCTTGCTCAG 900
 15 GGTGTCCCCG GTGACAACGT CGGCTTCAAC GTGAAGAAGC TTTCCGTCAA 950
 GGATATCCGC CGTGGTAACG TTGCCGGTGA CTCCAAGAAC GACCCCCCTG 1000
 CTGGCGCCGC CTCTTTCACC GCCCAGGTCA TCGTTCTCAA CCACCCCGGT 1050
 CAGGTCGGTG CCCGCTACGC CCCCGTCTC GACTGCCACA CTGCCCACAT 1100
 TGCCTGCAAG TTCGCCGAGC TCCTCGAGAA GATCGACCGC CGTACTGGTA 1150
 50 AGGCTGTTGA GGCCTCCCCC AAGTTCATCA AGTCTGGTGA TGCTGCCATC 1200
 GTCAAGATGA TTCCCTCCAA GCCCATGTGC GTTGAGGCTT TCACCGACTA 1250
 CCCTCCCCTC GGCCGTTTCG CCGTCCGTGA CATGCGTCAG ACCGTCGCCG 1300
 TCGGTGTCAT CAAGGCCGTC GACAAGTCCA CCGTGCCGC TGGCAAGGTC 1350

ACCAAGTCCG CTGCCAAGGC CGCCAAGAAG TAA

1383

5 2) INFORMATION FOR SEQ ID NO: 1995

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1383 bases
 (B) TYPE: Nucleic acid
 10 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Podospira anserina*
 (C) ACCESSION NUMBER: X74799

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1995

20 ATGGGCAAGG AGGACAAGAC TCACATCAAC GTCGTCGTTA TCGGCCACGT 50
 CGATTCCGGC AAGTCGACCA CCACTGGTCA CTTGATCTAC AAGTGCGGTG 100
 GTATTGACAA GCGTACCATC GAGAAGTTCG AGAAGGAAGC TGCTGAGCTC 150
 GGCAAGGGCT CTTTCAAGTA TGCCTGGGTT CTTGACAAGT TGAAGGCCGA 200
 25 GCGTGAGCGT GGTATCACCA TCGATATTGC CCTCTGGAAG TTCGAGACCC 250
 CCAAGTACTA TGTCACCGTC ATTGATGCCC CCGGCCATCG TGATTTTCATC 300
 AAGAACATGA TTAAGGTGAC TTCCAGGCC GATTGCGCCA TTCTCATCAT 350
 TGCCGCCGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA 400
 CCCGTGAGCA CGCTCTCCTC GCCTACACCC TCGGTGTGAA GCAGCTCATC 450
 30 GTCGCCATCA ACAAGATGGA CACCACCAAG TGGTCCGAGG CCCGCTTCAA 500
 CGAGATCATC AAGGAGACCT CCAACTTCAT CAAGAAGGTC GGCTACAACC 550
 CCAAGACTGT TGCCTTCGTC CCCATCTCCG GTTTCAACGG CGACAACATG 600
 CTTGAGGCTT CCACCAACTG CCCCTGGTAC AAGGGCTGGG AGAAGGAGGT 650
 CAAGGGTGGC AAGGCCACCG GCAAGACCCT CCTTGAGGCC ATCGACTCCA 700
 35 TCGAGCCCCC CAAGCGTCCC ACCGACAAGC CCCTCCGTCT TCCCCCTCCAG 750
 GATGTCTACA AGATCGGCGG TATCGGCACA GTCCCTGTCG GCCGTATCGA 800
 GACTGGTATC CTCAAGCCCG GTATGGTCGT TACCTTCGCT CCTTCCAACG 850
 TCACCACTGA AGTCAAGTCC GTCGAGATGC ACCACGAGCA GCTCGCTGAG 900
 GGTGTTCCCG GTGACAACGT TGGTTTCAAC GTGAAGAACG TCTCCGTCAA 950
 40 GGAAATCCGC CGTGGCAACG TTGCCGGTGA CTCCAAGAAC GACCCCCCCA 1000
 TGGGCGCCGC CTCTTTTCGAT GCCCAGGTCA TCGTCCTCAA CCACCCCGGC 1050
 CAGGTCGGTG CTGGTTACGC CCCCGTCTC GATTGCCACA CTGCCCACAT 1100
 CGCCTGCAAG TTCTCTGAGC TCCTGCAGAA GATCGACCGC CGTACTGGTA 1150
 AGGCCGTTGA GGAGAGCCCC AAGTTCATCA AGTCTGGTGA TGCTGCCATC 1200
 45 GTCAAGATGG TTCCCTCCAA GCCCATGTGC GTTGAGGCTT TCACTGAGTA 1250
 CCCTCCCCTC GGTCGTTTCG CCGTCCGTGA CATGCGTCAG ACCGTCGCTG 1300
 TCGGTGTCAT CAAGAAGGTC GAGAAGGCCG CTGCTGGTTC CGGCAAGGTT 1350
 ACCAAGTCCG CTGCCAAGGC TGGCAAGAAA TAA 1383

50

2) INFORMATION FOR SEQ ID NO: 1996

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1386 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 5 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:
 10 (A) ORGANISM: *Podospira curvicolla*
 (B) STRAIN: VLV
 (C) ACCESSION NUMBER: X96614

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1996

```

15 ATGGGCAAGG AGGACAAGAC TCACATCAAC GTCGTCGTTA TCGGCCACGT      50
   CGATTCCGGC AAGTCGACCA CCACTGGTCA CTTGATCTAC AAGTGCGGTG      100
   GTATTGACAA GCGTACCATC GAGAAGTTCG AGAAGGAAGC TGCTGAGCTC      150
   GGCAAGGGCT CTTTCAAGTA TGCCTGGGTT CTTGACAAGT TGAAGGCCGA      200
20 GCGTGAGCGT GGTATCACCA TTGATATCGC CCTCTGGAAG TTCGAGACCC      250
   CCAAGTACTA TGTCACCGTC ATCGATGCCC CCGGCCATCG TGATTTTCATC      300
   AAGAACATGA TTA CTGGTAC TTCCAGGCC GATTGCGCCA TTCTCATCAT      350
   TGCCGCCGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA      400
   CCCGTGAGCA CGCTCTCCTC GCCTACACCC TCGGTGTGAA GCAGCTCATC      450
25 GTCGCCATCA ACAAGATGGA CACCACCAA TGGTCCGAGG CCCGCTTCAA      500
   CGAGATCATC AAGGAGACCT CCAACTTCAT CAAGAAGGTC GGCTACAACC      550
   CCAAGACTGT TGCCTTCGTC CCCATCTCCG GTTTCAACGG CGACAACATG      600
   CTTGAGGCTT CCACCAACTG CCCCTGGTAC AAGGGTTGGG AGAAGGAGGT      650
   CAAGGGTGGC AAGGCTACTG GCAAGACCCT CCTCGAGGCC ATCGACTCCA      700
30 TCGAGCCCCC CAAGCGTCCC ACCGACAAGC CCCTCCGTCT TCCCCTTCAG      750
   GACGTTTACA AGATCGGCGG TATCGGCACA GTCCCTGTCTG GCCGTATCGA      800
   GACTGGTATC CTCAAGCCCG GTATGGTCGT TACCTTCGCC CCTTCCAACG      850
   TCACCACTGA AGTCAAGTCC GTCGAGATGC ACCACGAGCA GCTCTCTGAG      900
   GGTGTCCCCG GTGACAACGT TGGTTTCAAC GTGAAGAACG TCTCCGTCAA      950
35 GGAAATCCGC CGTGGCAACG TTGCCGGTGA CTCCAAGAAC GACCCCCCTC      1000
   TTGGCGCCGC TTCTTTTCGAT GCCCAGGTCA TCGTCTCTCAA CCACCCCGGC      1050
   CAGGTCGGTG CTGGTTACGC CCCCGTCTC GATTGCCACA CTGCCCACAT      1100
   CGCCTGCAAG TTCGCTGAGC TCCTGCAGAA GATCGATCGC CGTACTGGTA      1150
   AGGCTGTTGA GGAGAGCCCT AAGTTCATCA AGTCTGGTGA TGCTGCCATC      1200
40 GTCAAGATGA TTCCCTCCAA GCCCATGTGC GTTGAGGCTT TCACTGAGTA      1250
   CCCTCCCCTC GGTGCTTTCG CTGTCCGTGA CATGCGTCAG ACCGTCGCTG      1300
   TCGGTGTCAT CAAGAAGGTC GAGAAGGCCG CTGCTGGTTC CGGCAAGGTC      1350
   ACCAAGTCCG CTGCCAAGGC TGGTGGCAAG AAATAA      1386

```

45

2) INFORMATION FOR SEQ ID NO: 1997

- (i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 1383 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Sordaria macrospora*
 (B) STRAIN: 000
 (C) ACCESSION NUMBER: X96615

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1997

```

10 ATGGGTAAGG AAGACAAGGC TCACATCAAC GTCGTCGTTA TCGGCCACGT      50
   CGATTCCGGC AAGTCCACCA CTACCGGTCA CCTGATCTAC AAGTGCGGTG      100
   GTATCGACAA GCGTACCATC GAGAAGTTCG AGAAGGAAGC CGCTGAGCTC      150
   GGCAAGGGTT CCTTCAAGTA TGCCTGGGTT CTTGACAAGC TCAAGGCCGA      200
15 GCGTGAGCGT GGTATCACCA TCGATATCGC CCTCTGGAAG TTCGAGACTC      250
   CCAAGTACTA CGTCACCGTC ATCGATGCCC CCGGCCATCG TGATTTTCATC      300
   AAGAACATGA TCACTGGTAC CTCCCAGGCT GATTGCGCTA TTCTCATCAT      350
   TGCCGCTGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA      400
   CTCGTGAGCA CGCTCTTCTC GCCTACACCC TCGGTGTCAA GCAGCTCATC      450
20 GTTGCCATCA ACAAGATGGA CACCACCCAG TGGTCCCAGG CTCGTTTCGA      500
   GGAGATCATC AAGGAGACCA AGAACTTCAT CAAGAAGGTC GGCTACAACC      550
   CCGCCACCGT CGCTTTCGTC CCCATCTCCG GCTTCAACGG CGACAACATG      600
   CTTGAGGCCT CCACCAACTG CCCCTGGTAC AAGGGTTGGG AGAAGGAGAC      650
   CAAGGCCGGC AAGTCCACTG GCAAGACCCT CCTCGAGGCC ATCGACGCCA      700
25 TTGAGCAGCC CAAGCGCCCG ACCGACAAGC CCCTCCGTCT TCCCCTCCAG      750
   GATGTCTACA AGATCGGCGG TATCGGCACA GTGCCCGTCG GCCGTATCGA      800
   GACTGGTGTC CTCAAGCCCG GTATGGTCGT TACCTTCGCT CCTTCCAACG      850
   TCACCACTGA AGTCAAGTCC GTCGAGATGC ACCACGAGCA GCTTGCTCAG      900
   GGTGTTCCCG GTGACAACGT CGGCTTCAAC GTGAAGAACG TTTCCGTCAA      950
30 GGATATCCGT CGTGGTAACG TTGCCGGTGA CTCCAAGAAC GACCCCCCTG     1000
   TCGGCGCTGC CTCTTTCACC GCCCAGGTCA TCGTCCTTAA CCACCCCGGT     1050
   CAGGTCGGTG CCGGCTACGC TCCCGTCCTC GATTGCCACA CTGCCACAT      1100
   TGCCTGCAAG TTCGCCGAGC TCCTCGAGAA GATCGATCGC CGTACTGGTA      1150
   AGGCTGTTGA GACTTCTCCC AAGTTCATCA AGTCTGGTGA TGCTGCCATC      1200
35 GTCAAGATGA TTCCCTCCAA GCCCATGTGC GTCGAGGCTT TCACCGACTA      1250
   CCCTCCCCTC GGTCGTTTCG CCGTCCGTGA CATGCGTCAG ACCGTCGCTG      1300
   TCGGTGTCAT CAAGGCCGTC GACAAGACCC AGGCTGTCGC TGGCAAGGTC      1350
   ACCAAGTCTG CTGCCAAGGC TGCCAAGAAG TAA                          1383

```

10

2) INFORMATION FOR SEQ ID NO: 1998

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1383 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

50 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichoderma reesei*

(B) STRAIN: QM9414

(C) ACCESSION NUMBER: Z23012

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1998

5
 ATGGGTAAGG AGGACAAGAC TCACATCAAC GTGGTCGTCA TCGGCCACGT 50
 CGACTCCGGC AAGTCTACCA CCACTGGTCA CTTGATCTAC CAGTGCGGTG 100
 GTATCGACAA GCGTACCATT GAGAAGTTCG AGAAGGAAGC CGCCGAAGTC 150
 GGCAAGGGTT CCTTCAAGTA CGCGTGGGTT CTTGACAAGC TCAAGGCCGA 200
 10 GCGTGAGCGT GGTATCACCA TCGACATTGC CCTCTGGAAG TTCGAGACTC 250
 CCAAGTACTA TGTACCGTC ATTGACGCTC CCGGCCACCG TGACTTCATC 300
 AAGAACATGA TCACTGGTAC TTCCAGGCC GACTGCGCTA TCCTCATCAT 350
 CGCTGCCGGT ACTGGTGAGT TCGAGGCTGG TATCTCCAAG GATGGCCAGA 400
 CCCGTGAGCA CGCTCTGCTC GCCTACACCC TGGGTGTCAA GCAGCTCATC 450
 15 GTCGCCATCA ACAAGATGGA CACTGCCAAC TGGGCCGAGG CTCGTTACCA 500
 GGAAATCATC AAGGAGACTT CCAACTTCAT CAAGAAGGTC GGCTTCAACC 550
 CCAAGGCCGT TGCTTTCGTC CCCATCTCCG GCTTCAACGG TGACAACATG 600
 CTCACCCCTT CCACCAACTG CCCCTGGTAC AAGGGCTGGG AGAAGGAGAC 650
 CAAGGCTGGC AAGTTCACCG GCAAGACCCT CCTTGAGGCC ATCGACTCCA 700
 20 TCGAGCCCCC CAAGCGTCCC ACGGACAAGC CCCTGCGTCT TCCCCTCCAG 750
 GACGTCTACA AGATCGGTGG TATCGGAACA GTTCCCGTCG GCCGTATCGA 800
 GACTGGTGTC CTCAAGCCCG GTATGGTCGT TACCTTCGCT CCCTCCAACG 850
 TCACCACTGA AGTCAAGTCC GTCGAGATGC ACCACGAGCA GCTCGCTGAG 900
 GGCCAGCCTG GTGACAACGT TGGTTTCAAC GTGAAGAAGC TTTCCGTCAA 950
 25 GGAAATCCGC CGTGGCAACG TTGCCGGTGA CTCCAAGAAC GACCCCCCCA 1000
 TGGGCGCCGC TTCTTTCACC GCCCAGGTCA TCGTCATGAA CCACCCCGGC 1050
 CAGGTCGGTG CCGGCTACGC CCCCGTCCCT GACTGCCACA CTGCCCACAT 1100
 TGCCTGCAAG TTCGCCGAGC TCCTCGAGAA GATCGACCGC CGTACCGGTA 1150
 AGGCTACCGA GTCTGCCCCC AAGTTCATCA AGTCTGGTGA CTCCGCCATC 1200
 30 GTCAAGATGA TCCCCTCCAA GCCCATGTGC GTTGAGGCTT TCACCGACTA 1250
 CCCTCCCCTG GGTGTTTCG CCGTCCGTGA CATGCGCCAG ACCGTCGCTG 1300
 TCGGTGTCAT CAAGGCCGTC GAGAAGTCCT CTGCCGCCGC CGCCAAGGTC 1350
 ACCAAGTCCG CTGCCAAGGC CGCCAAGAAA TAA 1383

35

2) INFORMATION FOR SEQ ID NO: 1999

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 29 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

45 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1999

50 CATGTCAAYA TTGGTACTAT TGGTCATGT

29

2) INFORMATION FOR SEQ ID NO: 2000

1041

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 bases
 (B) TYPE: Nucleic acid
5 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2000
- CCACCYTCIC TCAMGTTGAA RCGTT 25
- 15 2) INFORMATION FOR SEQ ID NO: 2001
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 bases
20 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2001
- ACYACITTRA CIGCYGCIYAT YAC 23
- 30 2) INFORMATION FOR SEQ ID NO: 2002
- (i) SEQUENCE CHARACTERISTICS:
35 (A) LENGTH: 21 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2002
- CCIGARGARA GAGCIMGWGG T 21
- 45 2) INFORMATION FOR SEQ ID NO: 2003
- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2003

CATYTCRAIR TTGTCACCTG G

21

2) INFORMATION FOR SEQ ID NO: 2004

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1360 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida albicans*
- (B) STRAIN: SC5314
- (C) ACCESSION NUMBER: Genome project

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2004

| | | | | | |
|------------|------------|------------|------------|------------|------|
| GCTGCCTTCG | ACCGTTCTAA | ACCTCATGTC | AACATTGGTA | CTATTGGTCA | 50 |
| TGTTGATCAT | GGTAAACTA | CATTGACTGC | TGCTATCACC | AAAGTTTTAG | 100 |
| CCGAACAAGG | TGGTGCCAAC | TTCTTGGATT | ATGGTTCTAT | TGATAGAGCT | 150 |
| CCAGAAGAAA | GAGCTAGAGG | TATCACTATT | TCCACTGCCC | ACGTTGAATA | 200 |
| CGAAACCAAG | AACAGACACT | ATGCCCACGT | TGATTGTCCA | GGACACGCTG | 250 |
| ATTATATCAA | AAATATGATT | ACTGGTGCCG | CTCAAATGGA | TGGTGCTATC | 300 |
| ATTGTTGTTG | CTGCCACTGA | TGGTCAAATG | CCTCAAACCA | GAGAACATTT | 350 |
| GTTATTGGCC | AGACAAGTTG | GTGTTCAAGA | CTTGTTGTG | TTTGTCACAA | 400 |
| AAGTCGATAC | TATTGATGAC | CCTGAAATGT | TGGAATTAGT | CGAAATGGAA | 450 |
| ATGAGAGAAT | TGTTATCCAC | CTACGGTTTT | GATGGTGACA | ACACTCCAGT | 500 |
| TATTATGGGA | TCTGCTTTAA | TGGCTTTGGA | AGACAAGAAA | CCAGAAATTG | 550 |
| GTAAGGAAGC | TATCTTGAAA | TTGTTAGATG | CTGTCGATGA | ACACATTCCA | 600 |
| ACTCCATCAA | GAGACTTGGA | ACAACCATTT | TTGTTACCAG | TTGAAGACGT | 650 |
| GTTCTCCATC | TCCGGTAGAG | GAAGTGTGT | CACTGGTAGA | GTTGAAAGAG | 700 |
| GTGTTTTGAA | GAAGGGTGAA | GAAATCGAAA | TTGTTGGTGG | TTTTGACAAA | 750 |
| CCTTACAAGA | CTACTGTTAC | CGGTATTGAA | ATGTTCAAAA | AAGAATTAGA | 800 |
| CTCTGCTATG | GCTGGTGACA | ACTGTGGTGT | TTTGTTAAGA | GGTGTTAAAA | 850 |
| GAGATGAAAT | CAAGAGAGGT | ATGGTTTTGG | CCAAACCAGG | TACTGCTACT | 900 |
| TCTCACAAGA | AGTTCTTGGC | TTCTTGTAT | ATTTTGACTT | CCGAAGAAGG | 950 |
| TGGTCGTTCC | ACTCCATTTG | GTGAAGGTTA | CAAGCCTCAA | TGCTTCTTCA | 1000 |
| GAAGTAACGA | TGTCCTACC | ACATTTTCAT | TCCCAGAAGG | AGAAGGTGTT | 1050 |
| GATCATTCTC | AAATGATCAT | GCCAGGTGAC | AACATTGAAA | TGGTTGGTGA | 1100 |
| ATTGATCAAA | TCTTGTCAT | TAGAAGTCAA | CCAACGTTTC | AACTTGAGAG | 1150 |
| AAGGTGGTAA | AACTGTTGGT | ACTGGTTTGA | TTACCAGAAT | CATCGAATAA | 1200 |
| ACAGAATGTG | CACTGTGAAT | AATAAAAAGA | AAAGAGGTAT | ATATAGGTGA | 1250 |
| CTTTGTATTT | TGTATTGAAC | AATAAAATTC | TGTAAATAGT | AAGGGCCTCA | 1300 |
| GAAGTTTTGA | TTTGATTTAT | GCCATGTGGA | CTTGTAAGAG | TATCCTTCTC | 1350 |
| AAACTTCTTG | | | | | 1360 |

2) INFORMATION FOR SEQ ID NO: 2005

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1342 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Schizosaccharomyces pombe*

(C) ACCESSION NUMBER: Genome project

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2005

```

AAGCCGCATG TCAATATTGG TACTATTGGT CATGTTGACC ACGGTAAAAC      50
GACGTTGACG GCTGCTATTA CTAAATGCCT TTCTGATCTT GGTCAAGCTA     100
GTTTTATGGA TTATAGTCAA ATTGACAAGG CCCCCGAGGA AAAGGCACGT     150
GGTATTACCA TTTTATCTGC CCATGTTGAA TACGAACTG CTAATCGTCA     200
CTATGCCCAT GTGGATTGTC CTGGTCACGC CGATTACATT AAGAATATGA     250
TTACTGGTGC TGCTACAATG GATGGCGCTA TCATTGTTGT TTCTGCTACC     300
GATGGTCAAA TGCCTCAAAC TCGTGAACAT TTGCTTCTGG CTCGTCAAGT     350
CGGTGTAAAG CAAATTGTTG TATACATCAA TAAAGTCGAT ATGGTCGAGC     400
CTGATATGAT CGAGCTTGTC GAAATGGAAA TCGGTGAGCT ACTCTCCGAA     450
TACGGATTTG ATGGTGACAA TACTCCAATT GTTAGCGGCA GTGCTTTATG     500
TGCCTTAGAG GGTTCGTGAGC CTGAGATTGG TCTCAATAGT ATTACTAAAT     550
TGATGGAAGC TGTTGATAGT TATATTACTC TTCCTGAAAG AAAAACCGGAT     600
GTCCTTTTCT TGATGGCCAT CGAGGACGTT TTTTCAATTT CAGGTCGCGG     650
AACTGTAGTC ACTGGCCGTG TCGAGCGCGG TACTTTAAAG AAGGGTGCTG     700
AAATCGAAAT CGTCGGTTAT GGTAGCCATT TAAAGACTAC CGTTACTGGA     750
ATTGAAATGT TCAAAAAGCA GCTTGATGCC GCCGTTGCCG GTGACAATTG     800
TGGCCTTTTA CTTTCGTTCTA TCAAGCGAGA GCAATTAAAA CGTGGAATGA     850
TTGTCGCTCA ACCAGGAACC GTTGCTCCTC ATCAGAAATT CAAGGCATCA     900
TTCTATATTT TGACAAAAGA GGAAGGAGGT CGTCGTACCG GTTTCGTTGA     950
CAAGTATCGT CCCCACCTGT ACAGTCGTAC TTCCGACGTT ACTGTCGAAC    1000
TTACCCACCC TGATCCTAAC GACTCAGACA AAATGGTTAT GCCTGGAGAC    1050
AATGTCGAGA TGATCTGTAC GCTTATTCAC CCCATTGTCA TCGAAAAAGG    1100
ACAACGCTTC ACAGTTCGTG AGGGTGGAAG CACTGTAGGC ACAGCTTTGG    1150
TTACTGAACT TTTGGATTAG TGCATTATG AACTTATTGG CTTTAAAAAT    1200
TTTGCATGCT GAATACCAAT ATTATGTCCC TTCTCAGAAT TCTATAACTA    1250
CAGTGTCATT ATTGTAATAA GACTTTTGCA TCCATTGACA ATGGTATTTG    1300
ATACTTTTAT AGTTTCTACT ATTGTTAGCC AAAGTTATAA AA             1342

```

2) INFORMATION FOR SEQ ID NO: 2006

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2006

TGGAGCCGGT GAGCGTGG

18

2) INFORMATION FOR SEQ ID NO: 2007

1045

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2007

TGGAGCCAGT GAGCGTGG

18

2) INFORMATION FOR SEQ ID NO: 2008

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2008

TCTGGAGCCG ATGAGCGTG

19

2) INFORMATION FOR SEQ ID NO: 2009

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2009

CTGGAGCCAG TAAGCGTGG

19

2) INFORMATION FOR SEQ ID NO: 2010

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 861 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Double
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella pneumoniae*
 (B) STRAIN: KMK107
 (C) ACCESSION NUMBER: AF027199

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2010

```

ATGAGTATTC AACATTTTTCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
TTGCCTTCCT GTTTTGTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100
CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC      150
AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT      200
GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG      250
CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG      300
GTTAAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT      350
AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCTGCCA      400
ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG      450
CACAACTGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACC GGAGCT      500
GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGACG CCTGCAGCAA      550
TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT      600
TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC      650
ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG      700
GAGCCAGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT      750
GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC      800
TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA      850
AGCATTTGTA A                                     861

```

2) INFORMATION FOR SEQ ID NO: 2011

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 861 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella pneumoniae*
 (B) STRAIN: CLSiS L-491
 (C) ACCESSION NUMBER:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2011

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT      50
TTGCTTTCCT GTTTTGTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG      100
CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAGCT GGATCTCAAC      150
AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT      200
GAGCACTTTT AAAGTTCTGC TATGTGGTGC GGTATTATCC CGTGTTGACG      250
CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG      300
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT      350
AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA      400
ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG      450
CACAACTGG GGGATCATGT AACC CGCCTT GATCGTTGGG AACC GGAGCT      500
GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCAGCAA      550
TGGCAACAAC GTTGCGCAAA CTATTAAGTG GCGAACTACT TACTCTAGCT      600

```

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TCCCGGCAAC | AATTAATAGA | CTGGATGGAG | GCGGATAAAG | TTGCAGGACC | 650 |
| ACTTCTGCGC | TCGGCCCTTC | CGGCTGGCTG | GTTTATTGCT | GATAAATCTG | 700 |
| GAGCCAGTAA | GCGTGGATCT | CGCGGTATCA | TTGCAGCACT | GGGGCCAGAT | 750 |
| GGTAAGCCCT | CCCGTATCGT | AGTTATCTAC | ATGACGGGGA | GTCAGGCAAC | 800 |
| TATGGATGAA | CGAAATAGAC | AGATCGCTGA | GATAGGTGCC | TCACTGATTA | 850 |
| AGCATTGGTA | A | | | | 861 |

2) INFORMATION FOR SEQ ID NO: 2012

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2012

CCGCGGATTA TTAAACCGCC CTTCCGCGG-MR-HEG-ATGTCAGAGG GATAGATCCA 49

2) INFORMATION FOR SEQ ID NO: 2013

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 353 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Kluyvera ascorbata*
- (B) STRAIN: ATCC 33433

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2013

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| AGCTTAAGAA | CTCTTATCTG | GATTACGCGA | TGTCGGTCAT | TGTTGGCCGT | 50 |
| GCGCTGCCCG | ATGTCCGAGA | TGGCCTGAAG | CCGGTACACC | GTCGCGTACT | 100 |
| TTACGCCATG | AACGTATTGG | GCAATGACTG | GAACAAAGCC | TACAAAAAAT | 150 |
| CAGCCCGTGT | CGTGGGTGAC | GTGATCGGTA | AATATCACCC | GCATGGTGAT | 200 |
| ACTGCCGTCT | ATGACACTAT | CGTCCGTATG | GCACAGCCAT | TCTCACTGCG | 250 |
| ATACATGCTG | GATAGTGGTC | AAGGTAACCT | CGGTTCTGTC | GATGGCGACT | 300 |
| CCGCCGCAGC | GATGCGTTAT | ACGGAAATCC | GTATGTCGAA | AATCGCCCAT | 350 |
| GAG | | | | | 353 |

2) INFORMATION FOR SEQ ID NO: 2014

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 bases
- (B) TYPE: Nucleic acid

- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Kluyvera georgiana*
- (B) STRAIN: ATCC 51603

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2014

```

AGCTCCTATC TGGATTATGC GATGTCGGTC ATTGTTGGCC GTGCGCTGCC      50
AGATGTCCGA GATGGCCTGA AGCCGGTACA CCGTCGCGTA CTTTACGCCA      100
TGAACGTACT AGGCAATGAC TGGAACAAAG CCTATAAAAA ATCTGCCCGT      150
GTCGTTGGTG ACGTAATCGG TAAATACCAT CCCCATGGTG ACTCGGCGGT      200
CTATGACACG ATCGTCCGCA TGGCGCAGCC ATTCTCGCTG CGTTATATGC      250
TGGTAGACGG TCAGGGTAAC TTCGGTTCTA TCGACGGCGA CTCTGCGGCG      300
GCAATGCGTT ATACGGAAAT CCGTCTGGCG AAAATTGCCC ATGAACTGAT      350
GGCCG

```

2) INFORMATION FOR SEQ ID NO: 2015

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2015

CCAAGAAGCT CAAAAACATC TG

22

2) INFORMATION FOR SEQ ID NO: 2016

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2016

TADCCTGTCC AWACAGCCAT

20

2) INFORMATION FOR SEQ ID NO: 2017

(i) SEQUENCE CHARACTERISTICS:

1049

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2017

ACTTTGAATA AGGTCGGTCT AG

22

2) INFORMATION FOR SEQ ID NO: 2018

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2018

ACACTAAACA AGGTTGGTTT AG

22

2) INFORMATION FOR SEQ ID NO: 2019

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2019

ACACTAAACA AGGTCGGTCT AG

22

2) INFORMATION FOR SEQ ID NO: 2020

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2020

GTAGCTCCAG ATGAAATGTT TG

22

2) INFORMATION FOR SEQ ID NO: 2021

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2021

GTAGCTCCAG ACGAAATGTT TG

22

2) INFORMATION FOR SEQ ID NO: 2022

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2022

GTAGCTCCAG ATGAAACGTT TG

22

2) INFORMATION FOR SEQ ID NO: 2023

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2023

GTAAGTCCAG ATGAAATGTT TG

22

2) INFORMATION FOR SEQ ID NO: 2024

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid

1051

- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2024

AGTGAAAAGA TGGCTGCTGC

20

2) INFORMATION FOR SEQ ID NO: 2025

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2025

AGTGAGAAAA TGGCTGCTGC

20

2) INFORMATION FOR SEQ ID NO: 2026

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2026

TCCAAGCATG CATTATGCAA ACG

23

2) INFORMATION FOR SEQ ID NO: 2027

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2027

TCGGTCTAGA TAGAGCTAAA ACG

23

2) INFORMATION FOR SEQ ID NO: 2028

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2028

TATGCTCTTC AACATCACG

20

2) INFORMATION FOR SEQ ID NO: 2029

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2029

AGCCGTTGAG ACTTTGAATA AG

22

2) INFORMATION FOR SEQ ID NO: 2030

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2030

CTTAATGGTC TTGGTATCG

19

2) INFORMATION FOR SEQ ID NO: 2031

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1053

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2031

CGTGACTGGG GTTCTGCTAT GA 22

2) INFORMATION FOR SEQ ID NO:2032

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2032

CGTGACTGGG GATCATCAAT GA 22

2) INFORMATION FOR SEQ ID NO: 2033

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2033

CGTGACTGGG GTTCTGCCAT GA 22

2) INFORMATION FOR SEQ ID NO: 2034

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2034

ATCAAGAACA CTGGCTATGT AG 22

2) INFORMATION FOR SEQ ID NO: 2035

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2035

ATCAAGAACA CTGGCTACGT AG

22

2) INFORMATION FOR SEQ ID NO: 2036

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2036

ATCAAGAACA CTGGTTACGT AG

22

2) INFORMATION FOR SEQ ID NO: 2037

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2037

ATCAAAAATA CTGGTTATGT AG

22

2) INFORMATION FOR SEQ ID NO: 2038

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1055

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2038

ATCAAGAATA CTGGCTACGT AG

22

2) INFORMATION FOR SEQ ID NO: 2039

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2039

ATCAAAAACA CTGGCTATGT AG

22

2) INFORMATION FOR SEQ ID NO: 2040

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2040

TGTGACCCCA GACAAACCC

19

2) INFORMATION FOR SEQ ID NO: 2041

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2041

GTTGAGCGGC AGCACTATCT

20

2) INFORMATION FOR SEQ ID NO: 2042

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2042

CACGGGGATT TCTCTATTTA

20

2) INFORMATION FOR SEQ ID NO: 2043

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2043

CACGGGGATT ACTCTATTTA

20

2) INFORMATION FOR SEQ ID NO: 2044

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2044

ACCGTAAGTC GGCCAAGTCA

20

2) INFORMATION FOR SEQ ID NO: 2045

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2045

1057

GTTCTTTCTC CGTATCGTC

19

2) INFORMATION FOR SEQ ID NO: 2046

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2046

ACGGGGATTT TTCTATCTAT

20

2) INFORMATION FOR SEQ ID NO: 2047

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2160 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: CS109
- (C) ACCESSION NUMBER: Z49094

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2047

| | | | | | |
|-------------|------------|-------------|------------|------------|-----|
| ATGAACAAAC | CAACGATTCT | GCGCCTAATC | AAGTATCTGA | GCATTAGCTT | 50 |
| CTTAAGCTTG | GTTATCGCAG | CCATTGTCTT | AGGCGGAGGA | GTTTTTTTCT | 100 |
| ACTATGTTAG | CAAGGCTCCT | AGCCTATCCG | AGAGTAAACT | AGTTGCAACA | 150 |
| ACCTCTAGTA | AAATCTACGA | CAATAAAAAT | CAACTCATTG | CTGACTTGGG | 200 |
| TTCTGAACGC | CGCGTCAATG | CCCAAGCTAA | TGATATTCCC | ACAGATTTGG | 250 |
| TTAAGGCAAT | CGTTTCTATC | GAAGACCATC | GCTTCTTCGA | CCACAGGGGG | 300 |
| ATTGATACCA | TCCGTATCCT | GGGAGCTTTC | TTGCGCAATC | TGCAAAGCAA | 350 |
| TTCCCTCCAA | GGTGGATCAA | CTCTCACCCA | ACAGTTGATT | AAATTGACTT | 400 |
| ACTTTTCAAC | CTCGACTTCC | GACCAGACTA | TTTCTCGTAA | GGCTCAGGAA | 450 |
| GCTTGGGTTAG | CGATTCAAGT | AGAACA AAAA | GCAACCAAAC | AGGAAATCTT | 500 |
| GACCTACTAT | ATAAATAAGG | TCTACATGTC | TAATGGCAAC | TATGGAATGC | 550 |
| AGACAGCAGC | TCAAAACTAC | TATGGTAAAG | ACCTCAATAA | TTTAAGTTTA | 600 |
| CCTCAGTTAG | CCTTGCTGGC | TGGAATGCCT | CAGGCACCAA | ACCAATATGA | 650 |
| CCCCTATTCA | CATCCAGAAG | CAGCCCAAGA | CCGCCGAAAC | TTGGTCTTAT | 700 |
| CTGAAATGAA | AAATCAAGGT | TACATCTCTG | CTGAACAGTA | TGAGAAAGCA | 750 |
| GTCAATACAC | CAATTACTGA | TGGACTACAA | AGTCTCAAAT | CAGCAAGTAA | 800 |
| TTACCCTGCT | TACATGGATA | ATTACCTCAA | GGAAGTCATC | AATCAAGTTG | 850 |
| AAGAAGAAAC | TGGCTATAAC | CTTCTAACTA | CTGGGATGGA | TGTTTACACA | 900 |

1058

| | | | | | |
|-------------|------------|------------|-------------|------------|------|
| AATGTAGACC | AAGAAGCTCA | AAAACATCTG | TGGGATATCT | ACAACTCCGA | 950 |
| TCAATACGTC | TCTTACCCTG | ACGATGATTT | GCAAGTCGCA | TCTACGGTCG | 1000 |
| TAGATGTTTC | AAATGGTAAA | GTCATCGCAC | AACTTGGTGC | TCGTCATCAA | 1050 |
| GCAAGTAATG | TTTCATTCGG | TACCAACCAG | GCCGTAGAAA | CCAATCGTGA | 1100 |
| CTGGGGATCA | TCAATGAAAC | CAATCACTGA | CTATGCTCCC | GCTTTAGAAT | 1150 |
| ATGGAGTCTA | TGACTCTACT | GCTTCTATTG | TACATGATGT | CCCTTATAAC | 1200 |
| TATCCTGGCA | CTGATACTCC | ACTCTACAAC | TGGGATCATG | TCTACTTTGG | 1250 |
| AAACATTACA | ATCCAGTATG | CTCTTCAACA | ATCACGAAAT | GTCACAGCCG | 1300 |
| TTGAGACTTT | GAATAAGGTC | GGTCTAGATA | GAGCTAAAAC | CTTCCTTAAT | 1350 |
| GGTCTTG GTA | TCGACTATCC | AAGCATGCAT | TATGCAAACG | CCATTTCAAG | 1400 |
| TAACACAAC T | GAATCCAACA | AAAAATATGG | TGCAAGTAGT | GAAAAAATGG | 1450 |
| CTGCTGCCTA | CGCTGCTTTT | GCTAATGGTG | GTATTTATCA | CAAACCAATG | 1500 |
| TATATCAATA | AAATCGTCTT | TAGTGATGGT | AGCGAAAAAG | AATTTTCTGA | 1550 |
| TGCTGGTACA | CGAGCTATGA | AAGAGACTAC | TGCCCTATATG | ATGACTGAAA | 1600 |
| TGATGAAAAC | TGTTTTAACT | TACGGAACAG | GACGTGGAGC | CTACCTACCA | 1650 |
| TGGCTTCCAC | AAGCAGGTAA | GACAGGTACT | TCTAACTATA | CTGACGAAGA | 1700 |
| AATTGAAAAG | TATATCAAGA | ACACTGGTTA | CGTAGCTCCA | GATGAAATGT | 1750 |
| TTGTAGGGTA | TACCCGTAAA | TATGCAATGG | CTGTTTGGAC | AGGATACTCA | 1800 |
| AATCGTCTAA | CTCCAATCAT | CGGAGATGGT | TTCCCTTGTTG | CTGGTAAAGT | 1850 |
| CTATCGTTCA | ATGATAACTT | ACCTTTCTGA | AGATGACCAA | CCTGGAGATT | 1900 |
| GGACAATGCC | AGATGGCTTG | TATAGAAATG | GAGAATTCGT | ATTTAAAAAT | 1950 |
| GGTGCTCGTT | CTACGTGGAG | CTCACCTGCT | CCACAACAAC | CCCCATCAAC | 2000 |
| TGAAAGTTCA | AGCTCATCAT | CAGATAGTTC | AACTTCACAG | TCTAACTCAA | 2050 |
| CCACTCCAAG | CACAAATAAT | AGTACGACTA | CCAATCCTAA | CAATAATACG | 2100 |
| CAACAATCAA | ATACAACCCC | TGATCAACAA | AATCAGAATC | CTCAACCAGC | 2150 |
| ACAACCATAA | | | | | 2160 |

2) INFORMATION FOR SEQ ID NO: 2048

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2160 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: R6
- (C) ACCESSION NUMBER: M90527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2048

| | | | | | |
|-------------|-------------|-------------|------------|------------|-----|
| ATGAACAAAC | CAACGATTCT | GCGCCTAATC | AAGTATCTGA | GCATTAGCTT | 50 |
| CTTAAGCTTG | GTTATCGCAG | CCATTGTCTT | AGGCGGAGGA | GTTTTTTTCT | 100 |
| ACTACGTTAG | CAAGGCTCCT | AGCCTATCCG | AGAGTAAACT | AGTTGCAACA | 150 |
| ACTTCTAGTA | AAATCTACGA | CAATAAAAAT | CAACTCATTG | CTGACTTGGG | 200 |
| TTCTGAACGC | CGCGTCAATG | CCCAAGCTAA | TGATATTCCC | ACAGATTTGG | 250 |
| TTAAGGCAAT | CGTTTCTATC | GAAGACCATC | GCTTCTTCGA | CCACAGGGGG | 300 |
| ATTGATACCA | TCCGTATCCT | GGGAGCTTTC | TTGCGCAATC | TGCAAAGCAA | 350 |
| TTCCCTCCAA | GGTGGATCAG | CTCTCACTCA | ACAGTTGATT | AAGTTGACTT | 400 |
| ACTTTTCAAC | TTCGACTTCC | GACCAGACTA | TTTCTCGTAA | GGCTCAGGAA | 450 |
| GCTTG GTTAG | CGATT CAGTT | AGAACA AAAA | GCAACCAAGC | AAGAAATCTT | 500 |
| GACCTACTAT | ATAAATAAGG | TCTACATGTC | TAA TGGAAC | TATGGAATGC | 550 |

| | | | | | |
|------------|------------|-------------|------------|------------|------|
| AGACAGCAGC | TCAAAACTAC | TATGGTAAAG | ACCTCAATAA | TTTAAGTTTA | 600 |
| CCTCAGTTAG | CCTTGCTGGC | TGGAATGCCT | CAGGCACCAA | ACCAATATGA | 650 |
| CCCCTATTCA | CATCCAGAAG | CAGCCCCAAGA | CCGCCGAAAC | TTGGTCTTAT | 700 |
| CTGAAATGAA | AAATCAAGGC | TACATCTCTG | CTGAACAGTA | TGAGAAAGCA | 750 |
| GTCAATACAC | CAATTACTGA | TGGGCTACAA | AGTCTCAAAT | CAGCAAGTAA | 800 |
| TTACCCTGCT | TACATGGATA | ATTACCTCAA | GGAAGTCATC | AATCAAGTTG | 850 |
| AAGAAGAAAC | AGGCTATAAC | CTACTCACAA | CTGGGATGGA | TGTCTACACA | 900 |
| AATGTAGACC | AAGAAGCTCA | AAAACATCTG | TGGGATATTT | ACAATACAGA | 950 |
| CGAATACGTT | GCCTATCCAG | ACGATGAATT | GCAAGTCGCT | TCTACCATTG | 1000 |
| TTGATGTTTC | TAACGGTAAA | GTCATTGCCC | AGCTAGGAGC | ACGCCATCAG | 1050 |
| TCAAGTAATG | TTTCCTTCGG | AATTAACCAA | GCAGTAGAAA | CAAACCGCGA | 1100 |
| CTGGGGATCA | ACTATGAAAC | CGATCACAGA | CTATGCTCCT | GCCTTGGAGT | 1150 |
| ACGGTGTCTA | CGAGTCAACT | GCCACTATCG | TTCACGATGA | GCCCTATAAC | 1200 |
| TACCCTGGGA | CAAATACCCC | TGTTTATAAC | TGGGATAGGG | GCTACTTTGG | 1250 |
| CAACATCACC | TTGCAATACG | CCCTGCAACA | ATCGCGAAAC | GTCCAGCCG | 1300 |
| TGGAAACTCT | AAACAAGGTC | GGACTCAACC | GCGCCAAGAC | TTTCCTAAAT | 1350 |
| GGTCTCGGAA | TCGACTACCC | AAGTATTAC | TACTCAAATG | CCATTTCAAG | 1400 |
| TAACACAACC | GAATCAGACA | AAAAATATGG | AGCAAGTAGT | GAAAAGATGG | 1450 |
| CTGCTGCTTA | CGCTGCCTTT | GCAAATGGTG | GAAGTTACTA | TAAACCAATG | 1500 |
| TATATCCATA | AAGTCGTCTT | TAGTGATGGG | AGTGAAAAAG | AGTTCTCTAA | 1550 |
| TGTCGGAAC | CGTGCCATGA | AAGAAACGAC | AGCCTATATG | ATGACCGACA | 1600 |
| TGATGAAAAC | AGTCTTGAGT | TATGGAAC | GACGAAATGC | CTATCTTGCT | 1650 |
| TGGCTCCCTC | AGGCTGGTAA | AACAGGAACC | TCTAACTATA | CAGACGAGGA | 1700 |
| AATTGAAAAC | CACATCAAGA | CCTCTCAATT | TGTAGCACCT | GATGAACTAT | 1750 |
| TTGCTGGCTA | TACGCGTAAA | TATTCAATGG | CTGTATGGAC | AGGCTATTCT | 1800 |
| AACCGTCTGA | CACCACTTGT | AGGCAATGGC | CTTACGGTCG | CTGCCAAAGT | 1850 |
| TTACCGCTCT | ATGATGACCT | ACCTGTCTGA | AGGAAGCAAT | CCAGAGGATT | 1900 |
| GGAATATACC | AGAGGGGCTC | TACAGAAATG | GAGAATTCGT | ATTTAAAAAT | 1950 |
| GGTGCTCGTT | CTACGTGGAG | CTCACCTGCT | CCACAACAAC | CCCCATCAAC | 2000 |
| TGAAAGTTCA | AGCTCATCAT | CAGATAGTTC | AAGTTTACAG | TCTAGCTCAA | 2050 |
| CCACTCCAAG | CACAAATAAT | AGTACGACTA | CCAATCCTAA | CAATAATACG | 2100 |
| CAACAATCAA | ATACAACCCC | TGATCAACAA | AATCAGAATC | CTCAACCAGC | 2150 |
| ACAACCATAA | | | | | 2160 |

2) INFORMATION FOR SEQ ID NO: 2049

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2160 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: URU-E159
- (C) ACCESSION NUMBER: AF139890

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2049

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGAACAAAC | CAACGATTCT | GCGCCTAATC | AAGTATCTGA | GCATTAGCTT | 50 |
| CTTAAGCTTG | GTTATCGCAG | CCATTGTCTT | AGGCGGAGGA | GTTTTTTTCT | 100 |
| ACTACGTTAG | CAAGGCTCCT | AGCCTATCCG | AGAGTAAACT | AGTTGCAACA | 150 |
| ACTTCTAGTA | AAATCTACGA | CAATAAAAAT | CAACTCATTG | CTGACTTGGG | 200 |

| | | | | | |
|------------|------------|-------------|-------------|-------------|------|
| TTCTGAACGC | CGCGTCAATG | CCCAAGCTAA | TGATATTCCC | ACAGATTTGG | 250 |
| TTAAGGCAAT | CGTTTCTATC | GAAGACCATC | GCTTCTTCGA | CCACAGGGGG | 300 |
| ATTGATTCCA | TCCGTATCCT | GGGAGCTTTC | TTGCGCAATC | TGCAAAGTAA | 350 |
| TTCCCTCCAA | GGTGGATCAA | CTCTCACCCA | ACAGTTGATT | AAGTTGACTT | 400 |
| ACTTTTCAAC | CTCGACTTCC | GACCAGACTA | TTTCTCGTAA | GGCTCAGGAA | 450 |
| GCTTGGTTAG | CGATTCAAGT | AGAACAAAAA | GCAACCAAAC | AGGAAATCTT | 500 |
| GACCTACTAT | ATAAATAAGG | TCTACATGTC | TAATGGCAAC | TATGGAATGC | 550 |
| AGACAGCAGC | TCAAAACTAC | TATGGTAAAG | ACCTCAATAA | TTTAAGTTTA | 600 |
| CCTCAGTTAG | CCTTGCTGGC | TGGAATGCCT | CAGGCACCAA | ACCAATATGA | 650 |
| CCCCTATTCA | CATCCAGAAG | CAGCCCAAGA | CCGCCGAAAC | TTGGTCTTAT | 700 |
| CTGAAATGAA | AAATCAAGGT | TACATCTCTG | CTGAACAGTA | TGAGAAAGCA | 750 |
| GTCAATACAC | CAATTACTGA | TGGACTACAA | AGTCTCAAAT | CAGCAAGTAA | 800 |
| TTACCCTGCT | TACATGGATA | ATTACCTCAA | GGAGGTCATC | AATCAAGTAG | 850 |
| AACAAGAAAC | TGGCTATAAC | CTTCTAACTA | CTGGGATGGA | TGTTTACACA | 900 |
| AATGTAGACC | AAGAAGCTCA | AAAACATCTG | TGGGATATCT | ACAACTCCGA | 950 |
| TCAATACGTC | TCTTACCCTG | ACGATGATTT | GCAAGTCGCA | TCTACGGTCG | 1000 |
| TAGATGTTTC | AAATGGTAAA | GTCATCGCCC | AACTTGGAGC | TCGTCACCAA | 1050 |
| GCAAGTAACG | TTTCATTGCG | TACCAACCAA | GCTGTGGAAA | CCAATCGTGA | 1100 |
| CTGGGGTTCT | GCTATGAAAC | CAATCACCGA | TTATGCACCT | GCCATAGAAT | 1150 |
| ACGGTGTTTA | TGATTCCACT | GCAACTATGG | TTAATGATAT | TCCTTATAAC | 1200 |
| TATCCGGGAA | CAAGCACACC | TGTCTACAAC | TGGGATAGAG | CATATTTCTG | 1250 |
| TAATATTACT | CTGCAATATG | CTCTTCAACA | ATCACGAAAT | GTCACAGCCG | 1300 |
| TTGAGACTTT | GAATAAGGTC | GGTCTAGATA | GAGCTAAAAC | CTTCCTTAAT | 1350 |
| GGTCTTGCTA | TCGACTATCC | AAGCATGCAT | TATGCAAACG | CCATTTCAAG | 1400 |
| TAATACAACA | GAATCTAATA | AACAATACGG | AGCAAGTAGT | GAAAAAATGG | 1450 |
| CTGCTGCTTA | TGCTGCCTTT | GCAAATGGTG | GCACTTACTA | TAAACCAATG | 1500 |
| TATATCCATA | AAGTCGTCTT | CAGTGATGGA | AGTAAAAAAG | AGTTCTCTAA | 1550 |
| TGTCGGAACT | CGTGCCATGA | AGGAAACGAC | AGCCTATATG | ATGACCGACA | 1600 |
| TGATGAAAAC | AGTCTTGACT | TATGGAAC TG | GGCGTGGAGC | CTATCTTCCT | 1650 |
| TGGCTTCCTC | AAGCTGGTAA | AACAGGAACC | TCTAACTATA | CAGATGAGGA | 1700 |
| AGTTGAAAAC | CACATCAAGA | ACACTGGCTA | TGTAGCTCCA | GATGAAATGT | 1750 |
| TTGTTGGTTA | TACTCGTAAG | TATTCTATGG | CTGTATGGAC | AGGTTATTCTG | 1800 |
| AATCGTTTAA | CTCCTATCGT | TGGAGATGGT | TTCTAGTTTG | CAGCTAAAGT | 1850 |
| TTATCGCTCA | ATGATAACGT | ATCTATCAGA | AGATACTCAT | CCAGAAGACT | 1900 |
| GGACGATGCC | AGACGGACTT | TTCAGAAATG | GAGAATTTCGT | ATTTAAAAAT | 1950 |
| GGTGCTCGTT | CTACGTGGAA | CTCACCTGCT | CCACAACAAC | CCCCATCAAC | 2000 |
| TGAAAGTTCA | AGCTCATCAT | CAGATAGTTC | AACTTCACAG | TCTAGCTCAA | 2050 |
| CCACTCCAAG | CACAAATAAT | AGTACGACTA | CCAATCCTAA | CAATAATACG | 2100 |
| CAACAATCAA | ATACAACCCC | TGATCAACAA | AATCAGAATC | CTCAACCAGC | 2150 |
| ACAACCATAA | | | | | 2160 |

2) INFORMATION FOR SEQ ID NO: 2050

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: 8303; 35193
- (C) ACCESSION NUMBER: AF046230

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2050

| | | | | | | |
|------------|------------|----|------------|-------------|------------|-----|
| CTGTGGGATA | TCTACAAC | TC | CGATCAATAC | GTCTCTTACC | CTGACGATGA | 50 |
| TTTGCAAGTC | GCATCTACGG | TC | GTAGATGT | TTCAAATGGT | AAAGTCATCG | 100 |
| CCCAACTTGG | AGCTCGTCAC | CA | AGCAAGTA | ACGTTTCATT | TGGTACCAAC | 150 |
| CAAGCTGTGG | AAACCAATCG | TG | ACTGGGGT | TCTGCTATGA | AACCAATCAC | 200 |
| CGATTATGCA | CCTGCCATAG | AA | TACGGTGT | TTATGATTCC | ACTGCAACTA | 250 |
| TGGTTAATGA | TATTCCTTAT | AA | TATCCGG | GAACAAGCAC | ACCTGTCTAC | 300 |
| AACTGGGATA | GAGCATATTT | CG | GAATATT | ACTCTGCAAT | ATGCTCTTCA | 350 |
| ACAATCACGA | AATGTCACAG | CC | GTTGAGAC | TTTGAATAAG | GTCGGTCTAG | 400 |
| ATAGAGCTAA | AACCTTCCTT | AA | TGGTCTTG | GTATCGACTA | TCCAAGCATG | 450 |
| CATTATGCAA | ACGCCATTTC | AA | GAATACA | ACAGAATCTA | ATAAACATA | 500 |
| CGGAGCAAGT | AGTGAAAAAA | TG | GCTGCTGC | TTATGCTGCC | TTTGCAAATG | 550 |
| GTGGCACTTA | CTATAAACCA | AT | GTATATCC | ATAAAGTCGT | CTTCAGTGAT | 600 |
| GGAAGTAAAA | AAGAGTTCTC | TA | ATGTCGGA | ACTCGTGCCA | TGAAGGAAAC | 650 |
| GACAGCCTAT | ATGATGACCG | AC | ATGATGAA | AACAGTCTTG | ACTTATGGAA | 700 |
| CTGGGCGTGG | AGCCTATCTT | CC | TGGCTTC | CTCAAGCTGG | TAAAACAGGA | 750 |
| ACCTCTAACT | ATACAGATGA | GG | AAGTTGAA | AACCACATCA | AGAACACTGG | 800 |
| CTATGTAGCT | CCAGATGAAA | TG | TTTGTGG | TTATACTCGT | AAGTATTCTA | 850 |
| TGGCTGTATG | GACAGGTTAT | TC | GAATCGTT | TAACCTCCTAT | CGTTGGAGAT | 900 |
| GGTTTCCTAG | TTGCAGCTAA | AG | TTTATCGC | | | 930 |

2) INFORMATION FOR SEQ ID NO: 2051

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: 63509; M11
- (C) ACCESSION NUMBER: AF046238

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2051

| | | | | | | |
|------------|-------------|----|----------|------------|------------|-----|
| CTGTGGGATA | TTTACAATAC | AG | ACGAATAC | GTTGCCTATC | CAGATGATGA | 50 |
| CATGCAAGTA | GCTTCGACTA | TT | GTAGATGT | ATCTAATGGT | AACGTTATTG | 100 |
| CACAACTTGG | TGCTCGTCAT | CA | AGCAAGTA | ATGTTTCATT | CGGCACCAAC | 150 |
| CAGGCTGTGG | AGACCAATCG | TG | ACTGGGGT | TCTTCTATGA | AACCAATCAC | 200 |
| TGACTATGCT | CCCGCTTTAG | AA | TATGGAGT | CTATGACTCT | ACTGCTTCTA | 250 |
| TTGTACATGA | TGTTCCCTTAT | AA | TATCCCTG | GCACTGATAC | TCCAGTCTAC | 300 |
| AACTGGGATC | ATGTCTACTT | TG | GAAACATT | ACAATCCAGT | ATGCTCTTCA | 350 |
| ACAATCACGA | AATGTCACAG | CC | GTTGAGAC | TTTGAATAAG | GTCGGTCTAG | 400 |
| ATAGAGCTAA | AACCTTCCTT | AA | TGGTCTTG | GTATCGACTA | TCCAAGCATG | 450 |
| CATTATGCAA | ACGCCATTTC | AA | GAACACA | ACTGAATCCA | ACAAACAGTA | 500 |
| CGGTGCAAGT | AGTGAAAAGA | TG | GCTGCTGC | TTATGCCGCC | TTTTCTAATG | 550 |
| GTGGTATTTA | CCACAAACCA | AT | GTATATCA | ATAAAATCGT | CTTCAGTGAT | 600 |
| GGTAGTGAAA | AAGAATTTTC | TG | ATGCTGGT | ACACGAGCTA | TGAAAGAAAC | 650 |
| TACTGCCTAT | ATGATGACCG | AA | ATGATGAA | AACGTTTTTA | GCTTACGGAA | 700 |
| CCGGACGTGG | TGCTTACCTA | CC | ATGGCTTC | CACAAGCAGG | TAAGACAGGT | 750 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ACTTCTAACT | ATACTGACGA | CGAAATTGAA | AAGTATATCA | AGAACACTGG | 800 |
| CTACGTAGCT | CCAGATGAAA | TGTTTGTGG | TTATACTCGT | AAGTATTCTA | 850 |
| TGGCTGTATG | GACTGGTTAC | TCAAATCGTT | TAAGTCCAAT | CGTAGGAGAT | 900 |
| GGTTTCCTAG | TTGCTGCTAA | GGTTTATCGC | | | 930 |

2) INFORMATION FOR SEQ ID NO: 2052

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1195 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: #22/HA5
- (C) ACCESSION NUMBER: AB006877

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2052

| | | | | | |
|------------|------------|------------|------------|------------|------|
| GTCTCTTACC | CTGACGATGA | TTTGCAAGTC | GCATCTACGG | TCGTAGATGT | 50 |
| TTCCAATGGT | AAAGTCATCG | CCCAACTTGG | AGCTCGTCAC | CAAGCAAGTA | 100 |
| ACGTTTCATT | TGGTACCAAC | CAAGCTGTGG | AAACCAATCG | TGACTGGGGT | 150 |
| TCAACAATGA | AACCAATCAC | CGATTATGCA | CCTGCCATAG | AATACGGTGT | 200 |
| ATATGATTCC | ACTGCAACTA | TGGTTAATGA | TATTCCTTAT | AACTATCCGG | 250 |
| GAACAAGCAC | ACCTGTCTAC | AACTGGGATC | GAGCATATTT | TGGTAATATT | 300 |
| ACTCTGCAAT | ATGCCCTTCA | ACAATCACGT | AATGTTACAG | CAGTTGAAAC | 350 |
| ATTAAACAAG | GTTGGTTTAG | ATCGAGCCAA | AACTTTCCTA | AATGGCTTAG | 400 |
| GTATCGATTA | CCCAAGCATA | CACTATGCTA | ACGCCATTTC | AAGTAATACA | 450 |
| ACAGAATCTA | ACAAACATTA | CGGTGCAAGT | AGTGAAAAAA | TGGCTGCTGC | 500 |
| TTATGCCGCC | TTTGCTAATG | GTGGTATTTA | CCACAAACCA | ATGTATATCA | 550 |
| ATAAAATCGT | CTTTAGTGAT | GGTAGTGAGA | AAGAATTTTC | TGATGCTGGT | 600 |
| ACACGAGCTA | TGAAAGAAAC | TACTGCCTAT | ATGATGACCG | AAATGATGAA | 650 |
| AACTGTTTTA | GCTTACGGAA | CCGGACGTGG | TGCTTACCTA | CCATGGCTTC | 700 |
| CACAAGCAGG | TAAGACAGGT | ACTTCTAACT | ATACTGATGA | CGAAATCGAA | 750 |
| AAACATATCA | AGAATACTGG | CTACGTAGCT | CCAGATGAAA | TGTTTGTAGG | 800 |
| CTATACTCGT | AAGTATTCTA | TGGCTGTATG | GACTGGTTAC | TCAAATCGTT | 850 |
| TAAGTCCAAT | TGTTGGAGAT | GGTTTCCTAG | TTGCTGCTAA | GGTTTATCGC | 900 |
| TCAATGATAT | CGTATCTATC | AGAAGATGAC | CAACCTGGAG | ATTGGACGAT | 950 |
| GCCAGACGGC | CTGTTCCGAA | ATGGAGAATT | CGTATTTAAA | AATGGTGCTC | 1000 |
| GTTCTACGTG | GAACTCACCT | GCTCCACAAC | AACCCCATC | AACTGAAAGT | 1050 |
| TCAAGCTCAT | CATCAGATAG | TTCAACTTCA | CAGTCTAGCT | CAACCACTCC | 1100 |
| AAGCACAAAT | AATAGTACGA | CTACCAATCC | TAACAATAAT | ACGCAACAAT | 1150 |
| CAAATACAAC | CCCTGATCAA | CAAAATCAGA | ATCCTCAACC | AGCAC | 1195 |

2) INFORMATION FOR SEQ ID NO: 2053

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: 17619

(C) ACCESSION NUMBER: AF046237

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2053

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CTGTGGGATA | TTTACAATAC | AGACGAATAC | GTTGCCTATC | CAGACGATGA | 50 |
| ATTGCAAGTC | GCTTCTACCA | TTGTTGATGT | TTCTAACGGT | AAAGTCATTG | 100 |
| CCCAGCTAGG | AGCACGCCAT | CAGTCAAGTA | ATGTTTCCTT | CGGAATTAAC | 150 |
| CAAGCAGTAG | AAACAAACCG | CGACTGGGGA | TCAACTATGA | AACCGATCAC | 200 |
| AGACTATGCT | CCTGCCTTGG | AGTACGGTGT | CTACGATTCA | ACTGCTACTA | 250 |
| TCGTTACGGA | TGAGCCCTAT | AACTACCCTG | GGACAAATAC | TCCTGTTTAT | 300 |
| AACTGGGATA | GGGGCTACTT | TGGCAACATC | ACCTTGCAAT | ACGCCCTGCA | 350 |
| ACAATCGCGA | AACGTCCCAG | CCGTGGAAAC | TCTAAACAAG | GTCGGACTCA | 400 |
| ACCGCGCCAA | GACTTTCCTA | AATGGTCTCG | GAATCGACTA | CCCAAGTATT | 450 |
| CACTACTCAA | ATGCCATTTC | AAGTAACACA | ACCGAATCAG | ACAAAAAATA | 500 |
| TGGAGCAAGT | AGTGAAAAGA | TGGCTGCTGC | TTACGCTGCC | TTTGCAAATG | 550 |
| GTGGAACCTA | CTATAAACCA | ATGTATATCC | ATAAAGTCGT | CTTTAGTGAT | 600 |
| GGGAGTGAAA | AAGAGTTCTC | TAATGTGCGA | ACTCGTGCCA | TGAAGGAAAC | 650 |
| GACAGCCTAT | ATGATGACCG | AAATGATGAA | AACAGTCTTG | AGTTATGGAA | 700 |
| CTGGACGAAA | TGCCTATCTT | GCTTGGCTTC | CTCAAGCTGG | TAAAACAGGT | 750 |
| ACCTCTAACT | ATACAGATGA | AGAAATTGAA | AAACACATCA | AAAACACTGG | 800 |
| CTATGTAGCT | CCAGATGAAA | CGTTTGTTGG | TTATACTCGT | AAGTATTCTA | 850 |
| TGGCTGTATG | GACAGGTTAC | ACAAACCGTC | TCACACCAAT | TGTGGGGGAT | 900 |
| GGCTTCACAG | TTGCTGCCAA | AGTTTACCGC | | | 930 |

2) INFORMATION FOR SEQ ID NO: 2054

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 306 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: R6

(C) ACCESSION NUMBER: Y07845

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2054

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CGTCGCATTC | TCTACGGAAT | GAATGAATTG | GGTGTGACCC | CAGACAAACC | 50 |
| CCATAAAAAA | TCTGCTCGTA | TTACAGGGGA | TGTCATGGGT | AAATACCACC | 100 |
| CACACGGGGA | TTCCTCTATT | TATGAAGCCA | TGGTCCGTAT | GGCTCAATGG | 150 |
| TGGAGCTACC | GTTACATGCT | TGTAGATGGT | CATGGGAATT | TTGGTTCCAT | 200 |
| GGATGGAGAT | AGTGCTGCCG | CTCAACGTTA | TACCGAGGCA | CGTATGAGCA | 250 |
| AGATTGCTCT | GGAAATGCTT | CGTGATATCA | ACAAAAATAC | AGTTGATTTC | 300 |
| GTTGAT | | | | | 306 |

2) INFORMATION FOR SEQ ID NO: 2055

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2472 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: 7785
- (C) ACCESSION NUMBER: Z67739

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2055

| | | | | | |
|-------------|-------------|-------------|-------------|-------------|------|
| ATGTCTAACA | TTCAAAACAT | GTCCCTGGAG | GACATCATGG | GAGAGCGCTT | 50 |
| TGGTCGCTAC | TCCAAGTACA | TTATTCAAGA | CCGGGCTTTG | CCAGATATTC | 100 |
| GTGATGGGTT | GAAGCCGGTT | CAGCGCCGTA | TTCTTTATTC | TATGAATAAG | 150 |
| GATAGCAATA | CTTTTGACAA | GAGCTACCGT | AAGTCGGCCA | AGTCAGTCGG | 200 |
| GAACATCATG | GGGAATTTCC | ACCCACACGG | GGATTCTTCT | ATCTATGATG | 250 |
| CCATGGTTTCG | TATGTCACAG | AACTGGAAAA | ATCGTGAGAT | TCTAGTTGAA | 300 |
| ATGCACGGTA | ATAACGGTTC | TATGGACGGA | GATCCTCCTG | CGGCTATGCG | 350 |
| TTATACTGAG | GCACGTTTGT | CTGAAATTGC | AGGCTACCTT | CTTCAGGATA | 400 |
| TCGAGAAAAA | GACAGTTCCT | TTTGCATGGA | ACTTTGACGA | TACGGAGAAA | 450 |
| GAACCAACGG | TCTTGCCAGC | AGCCTTTCCA | AACCTCTTGG | TCAATGGTTC | 500 |
| GACTGGGATT | TCGGCTGGTT | ATGCCACAGA | CATTCCTCCC | CATAATTTAG | 550 |
| CTGAGGTCAT | AGATGCTGCA | GTTTACATGA | TTGACCACCC | AACTGCAAAG | 600 |
| ATTGATAAAC | TCATGGAATT | CTTACCTGGA | CCAGACTTCC | CTACAGGGGC | 650 |
| TATTATTTCAG | GGTCGTGATG | AAATCAAGAA | AGCTTATGAG | ACTGGGAAAG | 700 |
| GGCGCGTGGT | TGTTTCGTTCC | AAGACTGAAA | TTGAAAAGCT | AAAAGGTGGT | 750 |
| AAGGAACAAA | TCGTTATTAC | TGAGATTCCCT | TATGAAATCA | ATAAGGCCAA | 800 |
| TCTAGTCAAG | AAAATCGATG | ATGTTTCGTGT | TAATAACAAG | GTAGCTGGGA | 850 |
| TTGCTGAGGT | TCGTGATGAG | TCTGACCGTG | ATGGTCTTCG | TATCGCTATC | 900 |
| GAACTTAAGA | AAGACGCTAA | TACTGAGCTT | GTTCTCAACT | ACTTATTTAA | 950 |
| GTACACCGAC | CTACAAATCA | ACTACAACCT | TAATATGGTG | GCGATTGACA | 1000 |
| ATTTACACCC | TCGTGAGGTT | GGGATTGTTC | CAATCCTGTC | TAGCTACATC | 1050 |
| GCTCACCGTC | GAGAAGTGAT | TTTGGCGCGT | TCACGCTTTG | ACAAAGAAAA | 1100 |
| GGCTGAGAAA | CGTCTCCATA | TCGTGCAAGG | TTTGATTTCGT | GTGATTTCTGA | 1150 |
| TTTTGGATGA | AGTCATTGCT | CTTATCCGTG | CTTCTGAGAA | TAAGGCGGAC | 1200 |
| GCCAAGGAAA | ACCTCAAAGT | TAGCTATGAT | TTTACGGAAG | AACAGGCTGA | 1250 |
| GGCTATCGTA | ACTTTGCAAC | TGTACCGTTT | GACCAATACC | GATGTGGTTG | 1300 |
| TCTTGCAGGA | AGAAGAAGCA | GAGCTTCGTG | AGAAGATTGC | TATGCTGGCG | 1350 |
| GCTATTATCG | GTGATGAAAG | GACTATGTAC | AATCTCATGA | AGAAAGAACT | 1400 |
| TCGTGAGGTC | AAGAAGAAAT | TTGCAACTCC | TCGTTTGAGT | TCTTTAGAAG | 1450 |
| ACACTGCGAA | AGCAATTGAG | ATTGATACAG | CTAGTCTTAT | CGCTGAGGAA | 1500 |
| GATACCTACG | TCAGCGTGAC | CAAGGCAGGT | TACATCAAGC | GTACCAGTCC | 1550 |
| ACGTTTCCTTT | GCGGCTTCCA | CCTTGGAAGA | AATTGGCAAG | CGTGATGATG | 1600 |
| ACCGTTTGAT | TTTTGTTCAA | TCTGCCAAGA | CAACCCAGCA | CCTCTTGATG | 1650 |
| TTCACAAGTC | TTGGAAATGT | CATCTACAGA | CCAATCCATG | AGTTGGCAGA | 1700 |
| TATTCGTTGG | AAGGACATCG | GAGAGCATCT | GAGCCAAACC | ATCACAAACT | 1750 |
| TTGAAACGAA | TGAAGAAATC | CTTTATGTGG | AAGTACTGGA | TCAGTTTGAC | 1800 |
| GATGCGACAA | CCTACTTTGC | AGTGACTCGC | CTTGGTCAAA | TCAAACGGGT | 1850 |

| | | | | | |
|------------|-------------|------------|------------|------------|------|
| AGAGCGAAAA | GAATTCACCTC | CATGGCGGAC | CTATAGATCT | AAGTCTGTCA | 1900 |
| AGTATGCTAA | GCTCAAAGAC | GATACAGATC | AGATTGTAGC | AGTGGCTCCG | 1950 |
| ATTAAACTAG | ATGATGTTGT | CTTGGTTAGT | CAAAATGGTT | ATGCCCTGCG | 2000 |
| TTTCAATATC | GAAGAGGTTT | CGGTTGTCGG | TGCTAAGGCA | GCAGGTGTCA | 2050 |
| AGGCTATGAA | TTTGAAAGAA | GATGATGTCC | TCCAATCTGG | CTTTATCTGT | 2100 |
| AATACTTCGT | CCTTCTACCT | CTTGACCCAG | CGTGGAAGCT | TGAAACGTGT | 2150 |
| TTCTATTGAG | GAAATTCTAG | CAACCAGCCG | TGCCAAACGA | GGATTACAAG | 2200 |
| TCTTGCGTGA | GTTGAAAAAC | AAACCGCATC | GTGTCTTCTT | GGCAGGAGCA | 2250 |
| GTTGCAGAGC | AAGGATTTGT | TGGCGATTTT | TTCAGTACGG | AAGTGGATGT | 2300 |
| GAACGACCAA | ACTCTGCTTG | TCCAATCCAA | TAAAGGAACA | ATCTATGAAA | 2350 |
| GCCGATTGCA | AGACTTGAAC | TTGTCAGAAC | GCACTAGCAA | TGGAAGCTTC | 2400 |
| ATTTCTGACA | CGATTTTCAGA | TGAAGAAGTT | TTTGACGCTT | ATCTTCAGGA | 2450 |
| AGTAGTTACT | GAAGATAAAT | AA | | | 2472 |

2) INFORMATION FOR SEQ ID NO: 2056

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1212 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2056

| | | | | | |
|------------|------------|------------|------------|------------|------|
| AAGCTCAAAA | ACATCTGTGG | GATATCTACA | ACTCCGATCA | ATACGTCTCT | 50 |
| TACCCTGACG | ATGATTGCA | AGTCGCATCT | ACGGTCGTAG | ATGTTTCAAA | 100 |
| TGGTAAAGTC | ATTGCCCAAC | TTGGAGCTCG | TCACCAAGCA | AGTAACGTTT | 150 |
| CATTTGGCAC | CAACCAAGCT | GTGGAAACCA | ATCGTGACTG | GGGTTCAACA | 200 |
| ATGAAACCAA | TCACAGACTA | TGCTCCTGCC | TTGGAATACG | GTGTCTACGA | 250 |
| TTCAACTGCT | ACTATCGTTC | ACGATGAGCC | CTATAACTAT | CCTGGGACAG | 300 |
| ATACCCCTGT | CTATAACTGG | GATAGGGGCT | ACTTTGGCAA | TATCACCTTG | 350 |
| CAATACGCCC | TGCAACAATC | GCGAAACGTC | CCAGCCGTGG | AAACACTAAA | 400 |
| CAAGGTCGGG | CTCAACCGCG | CCAAGACTTT | TCTAAATGGT | CTCGGAATCG | 450 |
| ACTACCCAAG | TATTCACTAC | TCAAATGCCA | TTTCAAGTAA | CACAACCGAG | 500 |
| TCAGACAAAA | AATATGGAGC | AAGTAGTGAA | AAGATGGCTG | CTGCTTACGC | 550 |
| TGCCTTTGCA | AATGGTGGA | CTTACTATAA | ACCAATGTAT | ATCCATAAAG | 600 |
| TCGTCTTTAG | TGATGGAAGT | GAAAAAGAGT | TCTCTAATGT | CGGAACTCGC | 650 |
| GCCATGAAAG | AAACGACTGC | TTACATGATG | ACAGAAATGA | TGAAAACAGT | 700 |
| CTTGACGTAC | GGAATTGGTC | GTGGTGCCTA | CCTGCCTTGG | CTTCCTCAAG | 750 |
| CTGGTAAAC | AGGTACTTCT | AACTATACTG | ACGAAGAAAT | TGAAAAGTAT | 800 |
| ATCAAGAACA | CTGGTTACGT | AACTCCAGAT | GAAATGTTTG | TAGGGTATAC | 850 |
| CCGTAAATAT | GCAATGGCTG | TATGGACAGG | CTATTCTAAC | CGTCTGACAC | 900 |
| CACTTGTAGG | CGATGGCCTT | ACGGTCGCTG | CCAAGGTTTA | CCGCTCTATG | 950 |
| ATGACCTACC | TGTCTGAAGG | AAGCAATCCA | GAGGATTGGA | ATATACCAGA | 1000 |
| GGGGCTCTAC | AGAAATGGAG | AATTCGTATT | TAAAAATGGT | GCTCGTTCTA | 1050 |
| CGTGGAGCTC | ACCTGCTCCA | CAACAACCCC | CATCAACTGA | AAGTTCAAGC | 1100 |
| TCATCATCAG | ATAGTTCAAC | TTCACAGTCT | AGCTCAACCA | CTCCAAGCAC | 1150 |
| AAATAATAGT | ACGACTACCA | ATCCTAACAA | TAATACGCAA | CAATCAAATA | 1200 |
| CAACCCCTGA | TC | | | | 1212 |

2) INFORMATION FOR SEQ ID NO: 2057

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1242 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-17

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2057

| | | | | | |
|------------|------------|------------|-------------|------------|------|
| ATGTAGACCA | AGAAGCTCAA | AAACATCTGT | GGGATATTTA | CAATACAGAC | 50 |
| GAATACGTTG | CCTATCCAGA | CGATGAATTG | CAAGTCGCTT | CTACCATTGT | 100 |
| TGATGTTTCT | AACGGGAAAG | TTATTGCTCA | GTTAGGTTCT | CGTCACCAAT | 150 |
| CAAGCAATGT | TTCTTCGGA | ATCAACCAAG | CTGTTGAAAC | CAACCGTGAC | 200 |
| TGGGGTTCTG | CCATGAAGCC | AATCACAGAC | TATGCTCCTG | CCTTAGAGTA | 250 |
| TGACATCTAC | GACTCAACTG | CTTCGATTGT | ACATGATGTT | CCTTATAACT | 300 |
| ATCCAGGTAC | TGATACTCCC | CTCTACAACT | GGGATAAAGT | CTACTTTGGA | 350 |
| AATATTACAA | TCCAGTATGC | ACTTCAACAG | TCACGTAATG | TCACAGCCGT | 400 |
| TGAGACTTTG | AATAAGGTCG | GTCTAGATAG | AGCTAAAACC | TTTCTTAATG | 450 |
| GTCTTGGTAT | CGACTATCCA | AGCATGCATT | ATGCAAACGC | CATTTCAGT | 500 |
| AATACGACTG | AGTCAAACAA | AAAGTACGGA | GCAAGTAGTG | AGAAAATGGC | 550 |
| TGCTGCTTAC | GCTGCTTTTG | CTAACGGTGG | TATCTACCAT | AAACCAATGT | 600 |
| ATATCAACAA | AATCGTCTTT | AGCGATGGTA | GCTCAAAAGA | ATACGCTGAT | 650 |
| CCTGGTACTC | GTGCCATGAA | AGAGACGACC | GCCTATATGA | TGACAGAAAT | 700 |
| GATGAAGACT | GTCTTGGCAT | ACGGAACGGG | TCGTGGTGCT | TATCTCCCTT | 750 |
| GGCTACCTCA | AGCTGGTAAG | ACTGGTACAT | CAAACATATAC | AGATGATGAA | 800 |
| ATTGAAAAC | ACATCAAAAA | TACTGGTTAT | GTAGCTCCAG | ACGAAATGTT | 850 |
| TGTTGGTTAT | ACTCGCAAAT | ATTCAATGGC | GGTATGGACA | GGTTACTCAA | 900 |
| ACCGCCTGAC | TCCTATCGTT | GGTGATGGCT | TCTATGTTGC | AGCTAAGGTT | 950 |
| TACCGTTCAA | TGATGACTTA | TCTGTCTGAG | GATAACAACC | CTGGCGACTG | 1000 |
| GACTATGCCA | GAAGGTCTCT | ATCGAAGTGG | TGAGTTCGTC | TTTAAAAAAG | 1050 |
| GTGCTCGTTC | TGCATGGACT | GCTCCTGCTC | CGCAACAGGC | CCCAACACCA | 1100 |
| GAAAGTTCGA | GCTCGACATC | AGAAAGTTCA | ACTTCACAGT | CAAGCTCAAC | 1150 |
| TACTCCAAGC | ACGAATAATA | GTGCAAACAA | TAATACCAAT | AACCAGCAAC | 1200 |
| CAAATACAAC | GCCTGGTCAA | CAAAACCAGA | ACCAAAATCA | GA | 1242 |

2) INFORMATION FOR SEQ ID NO: 2058

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1225 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
 (B) STRAIN: StrR-18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2058

| | | | | | |
|-------------|------------|------------|------------|------------|------|
| GAAGCTCAAA | AACATCTGTG | GGATATTTAC | AATACAGACG | AATACGTTGC | 50 |
| CTATCCAGAC | GATGAATTGC | AAGTCGCTTC | TACCATTGTT | GATGTTTCTA | 100 |
| ACGGGAAAAGT | TATTGCTCAG | TTAGGTTCTC | GTCACCAATC | AAGCAATGTT | 150 |
| TCCTTCGGAA | TCAACCAAGC | TGTTGAAACC | AACCGTGA | GGGGTTCTGC | 200 |
| CATGAAGCCA | ATCACAGACT | ATGCTCCTGC | CTTAGAGTAT | GACATCTACG | 250 |
| ACTCAACTGC | TTCGATTGTA | CATGATGTTT | CTTATAACTA | TCCAGGTACT | 300 |
| GATACTCCCC | TCTACAAGT | GGATAAAGTC | TACTTTGGAA | ATATTACAAT | 350 |
| CCAGTATGCA | CTTCAACAGT | CACGTAATGT | CACAGCCGTT | GAGACTTTGA | 400 |
| ATAAGGTCGG | TCTAGATAGA | GCTAAAACCT | TTCTTAATGG | TCTTGGTATC | 450 |
| GACTATCCAA | GCATGCATTA | TGCAAACGCC | ATTTCAAGTA | ATACGACTGA | 500 |
| GTCAAACAAA | AAGTACGGAG | CAAGTAGTGA | GAAAATGGCT | GCTGCTTACG | 550 |
| CTGCTTTTGC | TAACGGTGGT | ATCTACCATA | AACCAATGTA | TATCAACAAA | 600 |
| ATCGTCTTTA | GCGATGGTAG | CTCAAAAGAA | TACGCTGATC | CTGGTACTCG | 650 |
| TGCCATGAAA | GAGACGACCG | CCTATATGAT | GACAGAAATG | ATGAAGACTG | 700 |
| TCTTGGCATA | CGGAACGGGT | CGTGGTGCTT | ATCTCCCTTG | GCTACCTCAA | 750 |
| GCTGGTAAGA | CTGGTACATC | AAACTATACA | GATGATGAAA | TTGAAAATA | 800 |
| CATCAAAAAT | ACTGGTTATG | TAGCTCCAGA | CGAAATGTTT | GTTGGTTATA | 850 |
| CTCGCAAATA | TTCAATGGCG | GTATGGACAG | GTTACTCAA | CCGCCTGACT | 900 |
| CCTATCGTTG | GTGATGGCTT | CTATGTTGCA | GCTAAGGTTT | ACCGTTCAAT | 950 |
| GATGACTTAT | CTGTCTGAGG | ATAACAACCC | TGGCGACTGG | ACTATGCCAG | 1000 |
| AAGGTCTCTA | TGCAAGTGGT | GAGTTCGTCT | TTAAAAAAGG | TGCTCGTTCT | 1050 |
| GCATGGACTG | CTCCTGCTCC | GCAACAGGCC | CCAACACCAG | AAAGTTCGAG | 1100 |
| CTCGACATCA | GAAAGTTCAA | CTTCACAGTC | AAGCTCAACT | ACTCCAAGCA | 1150 |
| CGAATAATAG | TGCAACAAT | AATACCAATA | ACCAGCAACC | AAATACAACG | 1200 |
| CCTGGTCAAC | AAAACCAGAA | CCAAA | | | 1225 |

2) INFORMATION FOR SEQ ID NO: 2059

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 554 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
 (B) STRAIN: StrR-38

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2059

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TGGCTACTTT | GGAAACATTA | CAGTCCAGTA | TGCTCTTCAA | CAATCACGAA | 50 |
| ATGTCACAGC | CGTTGAGACT | TTGAATAAGG | TGGTCTAGA | TAGAGCTAAA | 100 |
| ACCTTCCTCA | ATGGACTTGG | TATCGATTAT | CCAAGCATGC | TTTATGCAAA | 150 |
| CGCCATTTC | AGTAACACAA | CTGAATCCAA | CAAAAAGTAC | GGAGCAAGTA | 200 |
| GTGAAAAAAT | GGCCGCTGCC | TACGCAGCTT | TTGCTAATGG | TGGTACTTAC | 250 |
| CACAAACCAA | TGTATATCAA | TAAAATCGTC | TTTAGTGATG | GTAGTGAAAA | 300 |
| AGAATTTTCT | GATGCCGGTA | CTCGGGCTAT | GAAAGAAACT | ACTGCCTATA | 350 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TGATGACCGA | AATGATGAAA | ACTGTCTTAT | TATACGGAAC | CGGACGTGGA | 400 |
| GCCTACCTAC | CTTGGCTTCC | ACAAGCAGGT | AAGACAGGTA | CTTCTAACTA | 450 |
| TACTGACGAA | GAAATTGAAA | AGTATATCAA | GAATGCTGGT | TACGTAGCTC | 500 |
| CAGATGAAAT | GTTTGTGGT | TATACCCGCA | AATATGCAAT | GGCTGTTTGG | 550 |
| ACAG | | | | | 554 |

2) INFORMATION FOR SEQ ID NO: 2060

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1249 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-57

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2060

| | | | | | |
|------------|------------|------------|------------|------------|------|
| CAACTGGGAT | GGATGTCTAC | ACAAATGTAG | ACCAAGAAGC | TCAAAAACAT | 50 |
| CTGTGGGATA | TTTACAATAC | AGACGAATAC | GTTGCCTATC | CAGACGATGA | 100 |
| ATTGCAAGTC | GCTTCTACCA | TTGTTGATGT | TTCTAACGGT | AAAGTCATTG | 150 |
| CCCAGCTAGG | AGCACGCCAT | CAGTCAAGTA | ATGTTTCCTT | CGGAATTAAC | 200 |
| CAAGCAGTAG | AAACAAACCG | CGACTGGGGA | TCAACTATGA | AACCGATCAC | 250 |
| AGACTATGCT | CCTGCCTTGG | AGTACGGTGT | CTACGATTCA | ACTGCTACTA | 300 |
| TCGTTACGA | TGAGCCCTAT | AACTACCCTG | GGACAAATAC | TCCTGTTTAT | 350 |
| AACTGGGATA | GGGGCTACTT | TGGCAACATC | ACCTTGCAAT | ACGCCCTGCA | 400 |
| ACAATCGCGA | AACGTCCCAG | CCGTGGAAAC | TCTAAACAAG | GTCGGACTCA | 450 |
| ACCGCGCCAA | GACTTTCCTA | AATGGTCTAG | GAATCGACTA | CCCAAGTATT | 500 |
| CACTACTCAA | ATGCCATTTC | AAGTAACACA | ACCGAATCAG | ACAAAAAATA | 550 |
| TGGAGCAAGT | AGTGAAAAGA | TGGCTGCTGC | TTACGCTGCC | TTTGCAAATG | 600 |
| GTGGAACCTA | CTATAAACCA | ATGTATATCC | ATAAAGTCGT | CTTTAGTGAT | 650 |
| GGGAGTGAAA | AAGAGTTCTC | TAATGTCGGA | ACTCGTGCCA | TGAAGGAAAC | 700 |
| GACAGCCTAT | ATGATGACCG | ACATGATGAA | AACAGTCTTG | ACTTATGGAA | 750 |
| CTGGACGAAA | TGCCTATCTT | GCTTGGCTCC | CTCAGGCTGG | TAAAACAGGA | 800 |
| ACCTCTAACT | ATACAGACGA | GGAAATTGAA | AACCACATCA | AGACCTCTCA | 850 |
| ATTTGTAGCA | CCTGATGAAC | TATTTGCTGG | CTATACGCGT | AAATATTCAA | 900 |
| TGGCTGTATG | GACAGGCTAT | TCTAACCGTC | TGACACCACT | TGTAGGCAAT | 950 |
| GGCCTTACGG | TCGCTGCCAA | AGTTTACCGC | TCTATGATGA | CCTACCTGTC | 1000 |
| TGAAGGAAGC | AATCCAGAAG | ATTGGAATAT | ACCAGAGGGG | CTCTACAGAA | 1050 |
| ATGGAGAATT | CGTATTTAAA | AATGGTGCTC | GTTCTACGTG | GAGCTCACCT | 1100 |
| GCTCCACAAC | AACCCCATC | AACTGAAAGT | TCAAGCTCAT | CATCAGATAG | 1150 |
| TTCAACTTCA | CAGTCTAGCT | CAACCACTCC | AAGCACAAAT | AATAGTACGA | 1200 |
| CTACCAATCC | TAACAATAAT | ACGCAACAAT | CAAATACAAC | CCCTGATCA | 1249 |

2) INFORMATION FOR SEQ ID NO: 2061

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 579 bases
- (B) TYPE: Nucleic acid

- (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
(B) STRAIN: StrR-60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2061

| | | | | | |
|------------|------------|-------------|------------|------------|-----|
| GAAGTTGAAC | AGACTGGTCA | CCAAGCTCCT | TCATATCCGC | CAACACTGCA | 50 |
| ATTTTCTTGC | CACCTTCATT | GGCTGGAATG | GCAGAGAAAG | TCTCTAAAAT | 100 |
| CAGTTTCATA | GCAGTTGGAT | TGGCATTATA | AACATCTGAC | AGGATATCTG | 150 |
| CTCCATTGGC | TGCTTTCTTC | CACCTCGGTAC | GGTTACGCGT | CAATTCAAGA | 200 |
| TGTTGGAAGG | CCAAACGAAT | TTGCTCCTCT | GAAACTCCTT | CTTGCAAGGC | 250 |
| AACATAGGAT | GCAATCATAG | CATTTGTCGC | ATTGTACTTG | CCAGTTACTG | 300 |
| GCAAATCAAG | GGCTTGCTCT | AAGAAATTGA | CCTTGAAGGT | CAGACTATCT | 350 |
| TTGCGCTCAA | CCAAGTCGGT | AATTCCCAGC | TCTGCTCCTT | GACCAAAACG | 400 |
| AACCACCTTT | TTATCAGTTG | GCAAATAGTC | CTCTACGATA | GGGTCAGCCG | 450 |
| GCGCTAAAAG | CAAGGAACCT | GAAGCCATTC | CGTCTGCAAT | TTGCATTTTT | 500 |
| CCTTTAGCAA | TCTCAGAACG | GTCTTTGAAA | AAGGCCAAAT | GAGCTTCTCC | 550 |
| AACCAAGGTC | ACGATGGCTG | TATGGACAG | | | 579 |

2) INFORMATION FOR SEQ ID NO: 2062

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1216 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
(B) STRAIN: StrR-63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2062

| | | | | | |
|------------|-------------|------------|------------|------------|-----|
| AAGCTCAAAA | ACATCTGTGG | GATATCTACA | ACTCCGATCA | ATACGTCTCT | 50 |
| TACCCTGACG | ATGATTTGCA | AGTCGCATCT | ACGGTCGTAG | ATGTTTCAAA | 100 |
| TGGTAAAGTC | ATCGCACAAAC | TTGGTGCTCG | TCATCAAGCA | AGTAATGTTT | 150 |
| CATTCGGTAC | CAACCAGGCC | GTAGAAACCA | ATCGTGACTG | GGGATCATCA | 200 |
| ATGAAACCAA | TCACTGACTA | TGCTCCCGCT | TTAGAATATG | GAGTCTATGA | 250 |
| CTCTACTGCT | TCTATTGTAC | ATGATGTCCC | TTATAACTAT | CCTGGCACTG | 300 |
| ATACTCCACT | CTACAACCTG | GATCATGTCT | ACTTTGGAAA | CATTACAATC | 350 |
| CAGTATGCTC | TTCAACAATC | ACGAAATGTC | ACAGCCGTTG | AGACTTTGAA | 400 |
| TAAGGTCGGT | CTAGATAGAG | CTAAAACCTT | CCTTAATGGT | CTTGGTATCG | 450 |
| ACTATCCAAG | CATGCATTAT | GCAAACGCCA | TTTCAAGTAA | CACAACTGAA | 500 |
| TCCAACAAAA | AATATGGTGC | AAGTAGTGAA | AAAATGGCTG | CTGCCTACGC | 550 |
| TGCTTTTGCT | AATGGTGTA | TTTATCACAA | ACCAATGTAT | ATCAATAAAA | 600 |
| TCGTCTTTAG | TGATGGTAGC | GAAAAAGAAT | TTTCTGATGC | TGGTACACGA | 650 |
| GCTATGAAAG | AGACTACTGC | CTATATGATG | ACTGAAATGA | TGAAAACGTG | 700 |
| TTTAACTTAC | GGAACAGGAC | GTGGAGCCTA | CCTACCATGG | CTTCCACAAG | 750 |

| | | | | | |
|------------|------------|------------|------------|------------|------|
| CAGGTAAGAC | AGGTACTTCT | AACTATACTG | ACGAAGAAAT | TGAAAAGTAT | 800 |
| ATCAAGAACA | CTGGTTACGT | AGCTCCAGAT | GAAATGTTTG | TAGGGTATAC | 850 |
| CCGTAAATAT | GCAATGGCTG | TTTGGACAGG | ATACTCAAAT | CGTCTAACTC | 900 |
| CAATCATCGG | AGATGGTTTC | CTTGTTGCTG | GTAAAGTCTA | TCGTTCAATG | 950 |
| ATAACTTACC | TTTCTGAAGA | TGACCAACCT | GGAGATTGGA | CAATGCCAGA | 1000 |
| TGGCTTGTAT | AGAAATGGAG | AATTCGTATT | TAAAAATGGT | GCTCGTTCTA | 1050 |
| CGTGGAGCTC | ACCTGCTCCA | CAACAACCCC | CATCAACTGA | AAGTTCAAGC | 1100 |
| TCATCATCAG | ATAGTTCAAC | TTCACAGTCT | AGCTCAACCA | CTCCAAGCAC | 1150 |
| AAATAATAGT | ACGACTACCA | ATCCTAACAA | TAATACGCAA | CAATCAAATA | 1200 |
| CAACCCCTGA | TCAACA | | | | 1216 |

2) INFORMATION FOR SEQ ID NO: 2063

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 810 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: ATCC 700673

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2063

| | | | | | |
|------------|-------------|-------------|-------------|------------|-----|
| AATACGTCTC | TTACCCCTGAC | GATGATTTCG | AAGTCGCATC | CACGGTCGTA | 50 |
| GATGTTTCAA | ATGGTAAAGT | CATCGCCCAA | CTTGGAGCTC | GTCACCAAGC | 100 |
| AAGTAACGTT | TCATTTGGTA | CCAACCAAGC | TGTGGAAACC | AATCGTGACT | 150 |
| GGGGTTCAAC | AATGAAACCC | ATCACCGATT | ATGCACCTGC | CATAGAATAC | 200 |
| GGTGTATATG | ATTCCACTGC | AACTATGGTT | AATGATATTC | CTTATAACTA | 250 |
| TCCGGGAACA | AGCACACCTG | TCTACAACCTG | GGATCGAGCA | TATTTTGGTA | 300 |
| ATATTACTCT | GCAATATGCC | CTTCAACAAT | CTCGTAACGT | ACCCGCCGTT | 350 |
| GAGACACTAA | ACAAGGTTGG | TTTAGATAGA | GCTAAAAGTT | TCCTAAATGG | 400 |
| TTTAGGAATC | GACTATCCTG | TAATGCACTA | TTCAAATGCT | ATTTCAAGTA | 450 |
| ATACTACCGA | ATCTAGTAAA | CAGTACGGGG | CAAGTAGTGA | AAAAATGGCC | 500 |
| ACTGCCTATG | CCGCATTCGC | AAACGGCGGT | ATTTACCACA | AACCAATGTA | 550 |
| CATCAATAAG | GTTGTCTTTA | GCGATGGTAG | CGAAAAAGAA | TTTTCTGACC | 600 |
| CTGGCACAAG | AGCCATGAAA | GAAACGACTG | CTTACATGAT | GACAGAGATG | 650 |
| ATGAAAACAG | TCTGGACTTA | CGGAACTGGT | CGCGGTGCCT | ACCTACCTTG | 700 |
| GCTTCCACAA | GCAGGTAAAA | CAGGTACTTC | TAACATATACT | GACGAAGAAA | 750 |
| TTGAAAAGTA | TATCAAGAAC | ACTGGTTACG | TAGCTCCAGA | TGAAATGTTT | 800 |
| GTAGGGTATA | | | | | 810 |

2) INFORMATION FOR SEQ ID NO: 2064

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 782 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
(B) STRAIN: ATCC 700678

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2064

```

TCTTACCCTG ACGATGATTT GCAAGTCGCA TCTACGGTCG TAGATGTTTC      50
AAATGGTAAA GTCATCGCCC AACTTGGAGC TCGTCACCAA GCAAGTAACG      100
TTTCATTTGG TACCAACCAA GCTGTGGAAA CCAATCGTGA CTGGGGATCA      150
ACTATGAAAC CAATCACAGA CTATGCTCCT GCCTTGGAGT ACGGTGTCTA      200
CGATTCAACT GCTACTATCG TTCACGATGA GCCCTATAAC TACCCTGGGA      250
CAGATATCCC TCTCTATAAC TGGGATCGAG CATATTTTCGG TAATATTACT      300
CTGCAATATG CCCTTCAACA ATCTCGTAAC GTACCTGCCG TTGAAACACT      350
AAACAAGGTC GGTCTAGATA AGGCTAAAC CTTCCTTAAT GGTCTTGGA      400
TCGACTATCC AAGCATGCAT TATGCAAACG CCATTTCAAG TAATACAAC      450
GAATCCAACA AAAAATATGG TGCAAGTAGT GAAAAAATAG CTACCGCCTA      500
TGCCGCATTC GCAAATGGTG GTATTTACCA CAAACCAATG TACATCAATA      550
AAGTTGTCTT TAGCGATGGT AGCGAAAAAG AATTTTCTGA CCCTGGCACA      600
AGAGCCATGA AAGAAACGAC TGCTTACATG ATGACAGAAA TGATGAAAAC      650
AGTCTGGACG TACGGAACGT GTCGTGGTGC CTACCTGCCT TGGCTTCCTC      700
AAGCTGGTAA AACAGGTACC TCTAACTATA CTGACGAAGA AATTGAAAAG      750
TATATCAAGA AACTGGTTA CGTAGCTCCA GA                          782

```

2) INFORMATION FOR SEQ ID NO: 2065

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2065

CCAGGACGTG GAGGCGATCA CA

22

2) INFORMATION FOR SEQ ID NO: 2066

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2066

CACCGACAGC GAGCCGATCA GA

22

2) INFORMATION FOR SEQ ID NO: 2067

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2067

AGCTGAGCCA ATTCATGG

18

2) INFORMATION FOR SEQ ID NO: 2068

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1068

ATTCATGGAC CAGAACAAC

19

2) INFORMATION FOR SEQ ID NO: 2069

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2069

CGCTGTCGGG GTTGACCC

18

2) INFORMATION FOR SEQ ID NO: 2070

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1073

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2070

GTTGACCCAC AAGCGCCG

18

2) INFORMATION FOR SEQ ID NO: 2071

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2071

CGACTGTCGG CGCTGGGG

18

2) INFORMATION FOR SEQ ID NO: 2072

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3534 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*
- (B) STRAIN: Rv
- (C) ACCESSION NUMBER: L27989

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2072

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| GTGCTGGAAG | GATGCATCTT | GGCAGATTCC | CGCCAGAGCA | AAACAGCCGC | 50 |
| TAGTCCTAGT | CCGAGTCGCC | CGCAAAGTTC | CTCGAATAAC | TCCGTACCCG | 100 |
| GAGCGCCAAA | CCGGGTCTCC | TTCGCTAAGC | TGCGCGAACC | ACTTGAGGTT | 150 |
| CCGGGACTCC | TTGACGTCCA | GACCGATTCC | TTCGAGTGCC | TGATCGGTTC | 200 |
| GCCGCGCTGG | CGCGAATCCG | CCGCCGAGCG | GGGTGATGTC | AACCCAGTGG | 250 |
| GTGGCCTGGA | AGAGGTGCTC | TACGAGCTGT | CTCCGATCGA | GGACTTCTCC | 300 |
| GGGTCGATGT | CGTTGTCTTT | CTCTGACCCT | CGTTTCGACG | ATGTCAAGGC | 350 |
| ACCCGTCGAC | GAGTGCAAAG | ACAAGGACAT | GACGTACGCG | GCTCCACTGT | 400 |
| TCGTCAACGC | CGAGTTCATC | AACAACAACA | CCGGTGAGAT | CAAGAGTCAG | 450 |
| ACGGTGTTCA | TGGGTGACTT | CCCGATGATG | ACCGAGAAGG | GCACGTTCAT | 500 |
| CATCAACGGG | ACCGAGCGTG | TGGTGCTCAG | CCAGCTGGTG | CGGTCGCCCCG | 550 |
| GGGTGTACTT | CGACGAGACC | ATTGACAAGT | CCACCGACAA | GACGCTGCAC | 600 |
| AGCGTCAAGG | TGATCCCGAG | CCGCGGCGCG | TGGCTCGAGT | TTGACGTCGA | 650 |
| CAAGCGCGAC | ACCGTCGGCG | TGCGCATCGA | CCGCAAACGC | CGGCAACCGG | 700 |

| | | | | | |
|-------------|------------|------------|-------------|------------|------|
| TCACCGTGCT | GCTCAAGGCG | CTGGGCTGGA | CCAGCGAGCA | GATTGTCGAG | 750 |
| CGGTTCCGGT | TCTCCGAGAT | CATGCGATCG | ACGCTGGAGA | AGGACAACAC | 800 |
| CGTCGGCACC | GACGAGGCGC | TGTTGGACAT | CTACCGCAAG | CTGCGTCCGG | 850 |
| GCGAGCCCC | GACCAAAGAG | TCAGCGCAGA | CGCTGTTGGA | AAACTTGTTT | 900 |
| TTCAAGGAGA | AGCGCTACGA | CCTGGCCCCG | GTCGGTCGCT | ATAAGGTCAA | 950 |
| CAAGAAGCTC | GGGCTGCATG | TCGGCGAGCC | CATCACGTCG | TCGACGCTGA | 1000 |
| CCGAAGAAGA | CGTCGTGGCC | ACCATCGAAT | ATCTGGTCCG | CTTGACGAG | 1050 |
| GGTCAGACCA | CGATGACCGT | TCCGGGCGGC | GTCGAGGTGC | CGGTGGAAAC | 1100 |
| CGACGACATC | GACCACTTCG | GCAACCGCCG | CCTGCGTACG | GTCGGCGAGC | 1150 |
| TGATCCAAAA | CCAGATCCGG | GTCGGCATGT | CGCGGATGGA | GCGGGTGGTC | 1200 |
| CGGGAGCGGA | TGACCACCCA | GGACGTGGAG | GCGATCACAC | CGCAGACGTT | 1250 |
| GATCAACATC | CGGCCGGTGG | TCGCCGCGAT | CAAGGAGTTC | TTCGGCACCA | 1300 |
| GCCAGCTGAG | CCAATTCATG | GACCAGAACA | ACCCGCTGTC | GGGGTTGACC | 1350 |
| CACAAGCGCC | GACTGTCTGC | GCTGGGGCCC | GGCGGTCTGT | CACGTGAGCG | 1400 |
| TGCCGGGCTG | GAGGTCCGCG | ACGTGCACCC | GTCGCACTAC | GGCCGGATGT | 1450 |
| GCCCCGATCGA | AACCCCTGAG | GGGCCCAACA | TCGGTCTGAT | CGGCTCGCTG | 1500 |
| TCGGTGTACG | CGCGGGTCAA | CCCGTTCGGG | TTCATCGAAA | CGCCGTACCG | 1550 |
| CAAGGTGGTC | GACGGCGTGG | TTAGCGACGA | GATCGTGTAC | CTGACCGCCG | 1600 |
| ACGAGGAGGA | CCGCCACGTG | GTGGCACAGG | CCAATTCGCC | GATCGATCGC | 1650 |
| GACGGTCGCT | TCGTGAGGCC | GCGCGTGCTG | GTCCGCCGCA | AGGCGGGCGA | 1700 |
| GGTGGAGTAC | GTGCCCTCGT | CTGAGGTGGA | CTACATGGAC | GTCTCGCCCC | 1750 |
| GCCAGATGGT | GTCGGTGGCC | ACCGCGATGA | TTCCCTTCCT | GGAGCACGAC | 1800 |
| GACGCCAAC | GTGCCCTCAT | GGGGGCAAAC | ATGCAGCGCC | AGGCGGTGCC | 1850 |
| GCTGGTCCGT | AGCGAGGCCC | CGCTGGTGGG | CACCGGGATG | GAGCTGCGCG | 1900 |
| CGGCATCGA | CGCGGCGACG | TCGTGCTCGC | AAGAAAGCGG | CGTCATCGAG | 1950 |
| GAGGTGTGCG | CCGACTACAT | CACTGTGATG | CACGACAACG | GCACCCGGCG | 2000 |
| TACCTACCGG | ATGCGCAAGT | TTGCCCGGTC | CAACCACGGC | ACTTGCGCCA | 2050 |
| ACCAAGTGGC | CATCGTGGAC | GCGGGCGACC | GAGTCGAGGC | CGGTGAGGTG | 2100 |
| ATCGCCGACG | GTCCCTGTAC | TGACGACGGC | GAGATGGCGC | TGGGCAAGAA | 2150 |
| CCTGCTGGTG | GCCATCATGC | CGTGGGAGGG | CCACAACCTAC | GAGGACGCGA | 2200 |
| TCATCCTGTC | CAACCGCCTG | GTCGAAGAGG | ACGTGCTCAC | CTCGATCCAC | 2250 |
| ATCGAGGAGC | ATGAGATCGA | TGCTCGCGAC | ACCAAGCTGG | GTGCGGAGGA | 2300 |
| GATCACCCGC | GACATCCCGA | ACATCTCCGA | CGAGGTGCTC | GCCGACCTGG | 2350 |
| ATGAGCGGGG | CATCGTGCGC | ATCGGTGCCG | AGGTTCCGCA | CGGGGACATC | 2400 |
| CTGGTCCGCA | AGGTACCCCC | GAAGGGTGAG | ACCGAGCTGA | CGCCGGAGGA | 2450 |
| GCGGCTGCTG | CGTGCCATCT | TCGGTGAGAA | GGCCCGCGAG | GTGCGCGACA | 2500 |
| CTTCGCTGAA | GGTGCCGCAC | GGCGAATCCG | GCAAGGTGAT | CGGCATTCCG | 2550 |
| GTGTTTTCCC | GCGAGGACGA | GGACGAGTTG | CCGGCCGGTG | TCAACGAGCT | 2600 |
| GGTGCCTGTG | TATGTGGCTC | AGAAACGCAA | GATCTCCGAC | GGTGACAAGC | 2650 |
| TGGCCGGCCG | GCACGGCAAC | AAGGGCGTGA | TCGGCAAGAT | CCTGCCGGTT | 2700 |
| GAGGACATGC | CGTTCCCTTG | CGACGGCAC | CCGGTGGACA | TTATTTTGAA | 2750 |
| CACCCACGGC | GTGCCGCGAC | GGATGAACAT | CGGCCAGATT | TTGGGACCCC | 2800 |
| ACCTGGGTTG | GTGTGCCAC | AGCGCTGGA | AGGTCGACGC | CGCCAAGGGG | 2850 |
| GTTCCGGACT | GGGCCGCCAG | GCTGCCCGAC | GAAGTCTCG | AGGCGCATGC | 2900 |
| GAACGCCATT | GTGTGACGCG | CGGTGTTTGA | CGGCGCCGAG | GAGGCCGAGC | 2950 |
| TGCAGGGCCT | GTTGTGCTGC | ACGCTGCCCA | ACCGCGACGG | TGACGTGCTG | 3000 |
| GTCGACGCCG | ACGGCAAGGC | CATGCTCTTC | GACGGGCGCA | GCGGCGAGCC | 3050 |
| GTTCCCGTAC | CCGGTCACGG | TTGGCTACAT | GTACATCATG | AAGCTGCACC | 3100 |
| ACCTGGTGGA | CGACAAGATC | CACGCCCGCT | CCACCGGGCC | GTACTCGATG | 3150 |
| ATCACCCAGC | AGCCGCTGGG | CGGTAAGGCG | CAGTTCGGTG | GCCAGCGGTT | 3200 |
| CGGGGAGATG | GAGTGCTGGG | CCATGCAGGC | CTACGGTGCT | GCCTACACCC | 3250 |
| TGCAGGAGCT | GTTGACCATC | AAGTCCGATG | ACACCGTCGG | CCGCGTCAAG | 3300 |
| GTGTACGAGG | CGATCGTCAA | GGGTGAGAAC | ATCCCGGAGC | CGGGCATCCC | 3350 |
| CGAGTCGTTT | AAGGTGCTGC | TCAAAGAACT | GCACTGCTG | TGCCTCAACG | 3400 |
| TCGAGGTGCT | ATCGAGTGAC | GGTGCGGCGA | TCGAACTGCG | CGAAGGTGAG | 3450 |
| GACGAGGACC | TGGAGCGGGC | CGCGGCCAAC | CTGGGAATCA | ATCTGTCCCG | 3500 |
| CAACGAATCC | GCAAGTTTCG | AGGATCTTTC | GTAA | | 3534 |

2) INFORMATION FOR SEQ ID NO: 2073

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2073

CCGAGCAACA TGATTGAACC ATCCACCAAC TGGCTCGG

38

2) INFORMATION FOR SEQ ID NO: 2074

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2074

CCGAGCAACA TGATTGAAGC TTCCACCAAC TGGCTCGG

38

2) INFORMATION FOR SEQ ID NO: 2075

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2075

CCGAGCCAGG TTCTGAAGTC TCTGCATTAT TAGGTGCTCG G

41

2) INFORMATION FOR SEQ ID NO: 2076

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1076

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2076

CCGAGCYGAY AACATTTTCA GATTACCCCA RCGCTCGG 39

2) INFORMATION FOR SEQ ID NO: 2077

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2077

CCGAGCAACC GATCCAGCTC CAGCTACGCT CGG 33

2) INFORMATION FOR SEQ ID NO: 2078

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2078

CCGAGCCTTG GTCTTCGGCC AAATGAACGC TCGG 34

2) INFORMATION FOR SEQ ID NO: 2079

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2079

CCGAGCGTTC AGTTACTTCA GTCCAAGCCG GCTCGG 36

2) INFORMATION FOR SEQ ID NO: 2080

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2080

CCGAGCCGAA GAGGGCCAAG ATGTCGCTCG G

31

2) INFORMATION FOR SEQ ID NO: 2081

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2081

GRATYRTYAA AGTTGGTGAG GAAG

24

2) INFORMATION FOR SEQ ID NO: 2082

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2082

CMACTTCATC YCGCTTCGTA CC

22

2) INFORMATION FOR SEQ ID NO: 2083

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1078

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2083

CCGCCGATGT TCCGTAAATT ACTTGAIGAA GGTCGAGCCG GCGG 44

2) INFORMATION FOR SEQ ID NO: 2084

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2084

CACGCGTCAA CACCCGTACA AGTCGTCTTT TGC GCGTG 38

2) INFORMATION FOR SEQ ID NO: 2085

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2085

CAA ACTAAAG AACATATCTT GCTA 24

2) INFORMATION FOR SEQ ID NO: 2086

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2086

ATATAATTTG CATCACCTTC AAG 23

2) INFORMATION FOR SEQ ID NO: 2087

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2087

TCAGCTCGTG GGATTAGGAG AG

22

2) INFORMATION FOR SEQ ID NO: 2088

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2088

AGGCTTCACG CTGTTAGGCT GA

22

2) INFORMATION FOR SEQ ID NO: 2089

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2089

ATGCTGAACT TATTGACCTT

20

2) INFORMATION FOR SEQ ID NO: 2090

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1080

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2090

CGTTACTGGA GTCGAAATG

19

2) INFORMATION FOR SEQ ID NO: 2091

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2091

CGCGACTTGA GATGGAAGTT AGTGAGCTTC TTGGTCGCG

39

2) INFORMATION FOR SEQ ID NO: 2092

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2092

CGCGACGAAA GAACTTCCTG AAGGTCGTGC AGGTCCAG

38

2) INFORMATION FOR SEQ ID NO: 2093

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2093

TGTTGGCAAT CGAAGACACC

20

2) INFORMATION FOR SEQ ID NO: 2094

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2094

TTCAATTTCT TGACCTACTT TCAA

24

2) INFORMATION FOR SEQ ID NO: 2095

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2095

CGGTCGGGTT GAACGTGG

18

2) INFORMATION FOR SEQ ID NO: 2096

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2096

CGCGACCGGT ACCACGGCCA GTAATCGTGT CGCG

34

2) INFORMATION FOR SEQ ID NO: 2097

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1185 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

1082

(A) ORGANISM: *Mycoplasma pneumoniae*
 (B) STRAIN: ATCC 29342
 (C) ACCESSION NUMBER: AE000019

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2097

```

ATGGCAAGAG AGAAATTTGA CCGATCTAAA CCCCACGTTA ATGTAGGTAC      50
TATTGGCCAC ATTGACCACG GTAAAACAAC TTAAACAGCA GCTATTTGTA      100
CTGTATTAGC AAAAGAAGGT AAATCAGCTG CTACTCGTTA CGACCAAATC      150
GATAAGGCTC CGGAAGAAAA AGCACGGGGA ATTACGATTA ACTCCGCTCA      200
CGTGGAGTAC TCCTCTGACA AGCGTCACTA TGCTCACGTT GACTGTCCAG      250
GACACGCTGA CTACATTAAG AACATGATTA CTGGTGCTGC ACAAATGGAT      300
GGTGCCATTC TAGTAGTTTC AGCAACTGAC AGTGTTATGC CCCAAACCCG      350
TGAACACATT TTGTTGGCCC GCCAAGTGGG TG TGCCACGC ATGGTAGTGT      400
TCCTAAACAA GTGTGACATT GCAACTGATG AAGAAGTGCA AGAGTTAGTA      450
GCAGAAAGAG TACGTGACTT ATTAAC TTCT TACGGCTTTG ATGGCAAGAA      500
CACCCCTATT ATTTATGGTT CTGCACTTAA AGCGCTTGAA GGTGATCCTA      550
AGTGGGAAGC TAAGATCCAT GATTTAATGA ATGCAGTTGA TGAATGGATT      600
CCAACCTCTG AACGTGAAGT GGACAAACCC TTCTTGTTGG CAATCGAAGA      650
CACCATGACG ATTACTGGCC GTGGTACCGT GGTTACCGGT CGGGTTGAAC      700
GTGGTGAATT GAAAGTAGGT CAAGAAATTG AAATCGTTGG TTTACGTCCA      750
ATCCGTAAAG CAGTTGTTAC CGGAATCGAA ATGTTCAAAA AGGAACTTGA      800
TTCAGCAATG GCTGGGGACA ACGCTGGGGT ATTACTCCGT GGTGTGGACC      850
GTAAAGAAGT GGAACGTGGT CAAGTGTTAG CTAAACCAGG TTCGATTAAA      900
CCGCACAAGA AATTTAAAGC GGAAATCTAT GCTTTAAAGA AGGAAGAAGG      950
TGGTCGTCAC ACCGGTTTCT TAAACGGTTA CCGTCCCCAA TTCTACTTCC     1000
GTACTACAGA CGTTACTGGT TCGATTTCCC TACCAGAAAA CACCGAAATG     1050
GTGCTACCAG GTGACAATAC CTCGATTACA GTTGAATAA TTGCACCAAT     1100
TGCTTGTGAA AAAGGTAGTA AGTTCTCCAT CCGTGAAGGT GGTCGAACGG     1150
TTGGTGCTGG TTCAGTCACG GAAGTGCTTG AATAG                          1185
  
```

2) INFORMATION FOR SEQ ID NO: 2098

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2098

CTGAGTCACA CCGACAAACG TC

22

2) INFORMATION FOR SEQ ID NO: 2099

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2099

CCAGGACTGA ACGGGATACG AA

22

2) INFORMATION FOR SEQ ID NO: 2100

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2100

GCGAGACGAT AGGTTGTC

18

2) INFORMATION FOR SEQ ID NO: 2101

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2609 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*
- (B) STRAIN: H37Rv
- (C) ACCESSION NUMBER: Z79701

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2101

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CAGCCCGCGA | GCGTAACCTG | GCTGCGATTT | CCGGCGCGGA | TTTTCGCAGT | 50 |
| GCGGTTACGC | TCGGAAAGCG | CGGGCCTCGC | CCACGCGGCG | GATGATGTCA | 100 |
| GCGGGGTGGT | CCTCGGCGAC | GACCCGGACC | ACGATCCACC | CGTAGCGGTG | 150 |
| CTGGACTTTC | TCGTGCCCGA | GGATGTCTTT | CCGGTAGTGG | TAGCGACTGG | 200 |
| TCAGATGGTG | GTGCGCGTCA | TACTCGGCCG | CGACCTTGAT | GTCTTGCCAG | 250 |
| CCCATATCCA | AATGGGCTTC | CGCCCAGCCC | CATTCGTTGC | GCACCGCGAT | 300 |
| CTGCGTCTGG | GGGCGCGGAA | AGCCGGCGCG | GATCAACAAC | AAGCGCAGCC | 350 |
| AGGTTTCCTT | GGGGGACTGG | GCACCGCCGT | CGACGAGGTC | CAGAGCGGCT | 400 |
| CTTGCGGCCT | TCATGCCACG | GCGGCCCCGA | TAGCGCTCGA | TCAGCGGCTC | 450 |
| GACGTCGGCC | ACCTTCAAAT | CGGTGGCCTG | TATCAGGGCG | TGCACGGCCG | 500 |
| CGACGGCGGG | GTCCAATGGA | AATCGACTGG | TCAGGTCGAG | CGCCGTTTCG | 550 |
| TCCGGTGTGG | TCACGCGCAT | GCCCTCGATG | ACGCAGATCT | CGTCGGGCTC | 600 |
| GATGCGCTCT | TCCCAGACTT | GCAGCCCCGG | GGCACGGCGG | CGGTTGGTGT | 650 |
| CGATGATCGC | GGCGGGAAGA | TCCGCGTCGA | TCCACTTGGC | GCCATGGAAG | 700 |
| GCAGAAGCCG | AGTAGCCGGC | CAGCACGCCG | CGGCGGCGCG | AGCGCAGCCA | 750 |
| CAGCGCTTTT | GCACGCAATT | GCGCGGTCAG | TTCCACACCC | TGCGGCACGT | 800 |

1084

| | | | | | |
|------------|------------|-------------|------------|------------|------|
| ACACGTCTTT | ATGTAGCGCG | ACATACCTGC | TGCGCAATTC | GTAGGGCGTC | 850 |
| AATACACCCG | CAGCCAGGGC | CTCGCTGCCC | AGAAAGGGAT | CCGTCATGGT | 900 |
| CGAAGTGTGC | TGAGTCACAC | CGACAAACGT | CACGAGCGTA | ACCCCACTGC | 950 |
| GAAAGTTCCC | GCCGGAAATC | GCAGCCACGT | TACGCTCGTG | GACATACCGA | 1000 |
| TTTCGGCCCC | GCCGCGGCGA | GACGATAGGT | TGTCGGGGTG | ACTGCCACAG | 1050 |
| CCACTGAAGG | GGCCAAACCC | CCATTCGTAT | CCCGTTCAGT | CCTGGTTACC | 1100 |
| GGAGGAAACC | GGGGGATCGG | GCTGGCGATC | GCACAGCGGC | TGGCTGCCGA | 1150 |
| CGGCCACAAG | GTGGCCGTCA | CCCACCGTGG | ATCCGGAGCG | CCAAAGGGGC | 1200 |
| TGTTTGGCGT | CGAATGTGAC | GTCACCGACA | GCGACGCCGT | CGATCGCGCC | 1250 |
| TTACGCGCGG | TAGAAGAGCA | CCAGGGTCCG | GTCGAGGTGC | TGGTGTTCAA | 1300 |
| CGCCGGCCTA | TCCGCGGACG | CATTCTCAT | GCGGATGACC | GAGGAAAAGT | 1350 |
| TCGAGAAGGT | CATCAACGCC | AACCTCACCG | GGGCGTTCCG | GGTGGCTCAA | 1400 |
| CGGGCATCGC | GCAGCATGCA | GCGCAACAAA | TTCGGTCGAA | TGATATTCAT | 1450 |
| AGGTTTCGGT | TCCGGCAGCT | GGGGCATCGG | CAACCAGGCC | AACTACGCAG | 1500 |
| CCTCCAAGGC | CGGAGTGATT | GGCATGGCCC | GCTCGATCGC | CCGCGAGCTG | 1550 |
| TCGAAGGCAA | ACGTGACCGC | GAATGTGGTG | GCCCCGGGCT | ACATCGACAC | 1600 |
| CGATATGACC | CGCGCGCTGG | ATGAGCGGAT | TCAGCAGGGG | GCGCTGCAAT | 1650 |
| TTATCCCAGC | GAAGCGGGTC | GGCACCCCCG | CCGAGGTCGC | CGGGGTGGTC | 1700 |
| AGCTTCCTGG | CTTCCGAGGA | TGCGAGCTAT | ATCTCCGGTG | CGGTCATCCC | 1750 |
| GGTCGACGGC | GGCATGGGTA | TGGGCCACTG | ACACAACACA | AGGACGCACA | 1800 |
| TGACAGGACT | GCTGGACGGC | AAACGGATTC | TGGTTAGCGG | AATCATCACC | 1850 |
| GACTCGTCGA | TCGCGTTTCA | CATCGCACGG | GTAGCCCAGG | AGCAGGGCGC | 1900 |
| CCAGCTGGTG | CTCACCGGGT | TCGACCGGCT | GCGGCTGATT | CAGCGCATCA | 1950 |
| CCGACCGGCT | GCCGGCAAAG | GCCCCGCTGC | TCGAACTCGA | CGTGCAAAAC | 2000 |
| GAGGAGCACC | TGGCCAGCTT | GGCCGGCCGG | GTGACCGAGG | CGATCGGGGC | 2050 |
| GGGCAACAAG | CTCGACGGGG | TGGTGCAATC | GATTGGGTTT | ATGCCGCAGA | 2100 |
| CCGGGATGGG | CATCAACCCG | TTCTTCGACG | CGCCCTACGC | GGATGTGTCC | 2150 |
| AAGGGCATCC | ACATCTCGGC | GTATTTCGTAT | GCTTCGATGG | CCAAGGCGCT | 2200 |
| GCTGCCGATC | ATGAACCCCG | GAGGTTCCAT | CGTCGGCATG | GACTTCGACC | 2250 |
| CGAGCCGGGC | GATGCCGGCC | TACAACTGGA | TGACGGTCGC | CAAGAGCGCG | 2300 |
| TTGGAGTCGG | TCAACAGGTT | CGTGGCGCGC | GAGGCCGGCA | AGTACGGTGT | 2350 |
| GCGTTTGAAT | CTCGTTGCCG | CAGGCCCTAT | CCGGACGCTG | GCGATGAGTG | 2400 |
| CGATCGTCGG | CGGTGCGCTC | GGCGAGGAGG | CCGGCGCCCA | GATCCAGCTG | 2450 |
| CTCGAGGAGG | GCTGGGATCA | GCGCGCTCCG | ATCGGCTGGA | ACATGAAGGA | 2500 |
| TGCGACGCCG | GTCGCCAAGA | CGGTGTGCGC | GCTGCTGTCT | GACTGGCTGC | 2550 |
| CGGCGACCAC | GGGTGACATC | ATCTACGCCG | ACGGCGGCGC | GCACACCCAA | 2600 |
| TTGCTCTAG | | | | | 2609 |

2) INFORMATION FOR SEQ ID NO: 2102

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2102

CACCTTCACC CTGACCGACG

20

2) INFORMATION FOR SEQ ID NO: 2103

1085

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2103

CGAACCAGCG GAAATAGTTG GAC

23

2) INFORMATION FOR SEQ ID NO: 2104

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2104

CTGGGCATGG CIGGAGTC

18

2) INFORMATION FOR SEQ ID NO: 2105

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3297 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*
- (B) STRAIN: H37rv
- (C) ACCESSION NUMBER: U68480

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2105

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGACACAGT | GCGCGAGCAG | ACGCAAAGC | ACCCCAAATC | GGGCGATTTT | 50 |
| GGGGGCTTTT | GCGTCTGCTC | GCGGGACGCG | CTGGGTGGCC | ACCATCGCCG | 100 |
| GGCTGATTGG | CTTTGTGTTG | TCGGTGGCGA | CGCCGCTGCT | GCCCGTCGTG | 150 |
| CAGACCACCG | CGATGCTCGA | CTGGCCACAG | CGGGGGCAAC | TGGGCAGCGT | 200 |
| GACCGCCCCG | CTGATCTCGC | TGACGCCGGT | CGACTTTACC | GCCACCGTGC | 250 |
| CGTGCGACGT | GGTGCGCGCC | ATGCCACCCG | CGGGCGGGGT | GGTGCTGGGC | 300 |
| ACCGCACCCA | AGCAAGGCAA | GGACGCCAAT | TTGCAGGCGT | TGTTCTGTCG | 350 |
| CGTCAGCGCC | CAGCGCGTGG | ACGTCACCGA | CCGCAACGTG | GTGATCTTGT | 400 |
| CCGTGCCGCG | CGAGCAGGTG | ACGTCCCCCG | AGTGTCACG | CATCGAGGTC | 450 |

1086

| | | | | | |
|-------------|-------------|------------|------------|-------------|------|
| ACCTCTACCC | ACGCCGGCAC | CTTCGCCAAC | TTCGTCGGGC | TCAAGGACCC | 500 |
| GTCGGGCGCG | CCGCTGCGCA | GCGGCTTCCC | CGACCCCAAC | CTGCGCCCGC | 550 |
| AGATTGTCGG | GGTGTTCACC | GACCTGACCG | GGCCCGCGCC | GCCCGGGCTG | 600 |
| GCGGTCTCGG | CGACCATCGA | CACCCGGTTC | TCCACCCGGC | CGACCACGCT | 650 |
| GAAACTGCTG | GCGATCATCG | GGGCGATCGT | GGCCACCGTC | GTCGCACTGA | 700 |
| TCGCGTTGTG | GCGCCTGGAC | CAGTTGGACG | GGCGGGGCTC | AATTGCCCAG | 750 |
| CTCCTCCTCA | GGCCGTTCCG | GCCTGCATCG | TGCGCGGGCG | GCATGCGCCG | 800 |
| GCTGATTCCG | GCAAGCTGGC | GCACCTTCAC | CCTGACCGAC | GCCGTGGTGA | 850 |
| TATTCGGCTT | CCTGCTCTGG | CATGTCATCG | GCGCGAATTC | GTCGGACGAC | 900 |
| GGCTACATCC | TGGGCATGGC | CCGAGTCGCC | GACCACGCCG | GCTACATGTC | 950 |
| CAACTATTTC | CGCTGGTTTC | GCAGCCCGGA | GGATCCCTTC | GGCTGGTATT | 1000 |
| ACAACCTGCT | GGCGCTGATG | ACCCATGTCA | GCGACGCCAG | TCTGTGGATG | 1050 |
| CGCCTGCCAG | ACCTGGCCGC | CGGGCTAGTG | TGCTGGCTGC | TGCTGTGCGG | 1100 |
| TGAGGTGCTG | CCCCGCCTCG | GGCCGGCGGT | GGAGGCCAGC | AAACCCGCCT | 1150 |
| ACTGGGCGGC | GGCCATGGTC | TTGCTGACCG | CGTGGATGCC | GTTCAACAAC | 1200 |
| GGCCTGCGGC | CGGAGGGCAT | CATCGCGCTC | GGCTCGCTGG | TCACCTATGT | 1250 |
| GCTGATCGAG | CGGTCCATGC | GGTACAGCCG | GCTCACACCG | GCGGCGCTGG | 1300 |
| CCGTCGTTAC | CGCCGCATTC | ACACTGGGTG | TGCAGCCAC | CGGCCTGATC | 1350 |
| GCGGTGGCCG | CGCTGGTGGC | CGGCGGCCGC | CCGATGCTGC | GGATCTTGTT | 1400 |
| GCGCCGYCAT | CGCCTGGTCG | GCACGTTGCC | GTTGGTGTGC | CCGATGCTGG | 1450 |
| CCGCCGGCAC | CGTCATCCTG | ACCGTGGTGT | TGCGCGACCA | GACCCGTGTCA | 1500 |
| ACGGTGTGTT | AAGCCACCAG | GGTTCGCGCC | AAAATCGGGC | CGAGCCAGGC | 1550 |
| GTGGTATACC | GAGAACCTGC | GTTACTACTA | CCTCATCCTG | CCCACCGTCG | 1600 |
| ACGGTTCGCT | GTGCGGCGCG | TTCGGCTTTT | TGATCACCGC | GCTATGCGCTG | 1650 |
| TTCACCGCGG | TGTTTCATCAT | GTTGCGGCGC | AAGCGAATTC | CCAGCGTGCC | 1700 |
| CCGCGGACCG | GCGTGGCGGC | TGATGGGCGT | CATCTTCGGC | ACCATGTTCT | 1750 |
| TCCTGATGTT | CACGCCCACC | AAGTGGGTGC | ACCACTTCGG | GCTGTTTCGCC | 1800 |
| GCCGTAGGGG | CGGCGATGGC | CGCGCTGACG | ACGGTGTGTT | TATCCCCATC | 1850 |
| GGTGCTGCGC | TGGTCGCGCA | ACCGGATGGC | GTTCTTGCGG | GCGTTATTCT | 1900 |
| TCCTGCTGGC | GTTGTGTTGG | GCCACCACCA | ACGGCTGGTG | GTATGTCTCC | 1950 |
| AGCTACGGTG | TGCCGTTCOA | CAGCGCGATG | CCGAAGATCG | ACGGGATCAC | 2000 |
| AGTCAGCACA | ATCTTTTTCG | CCCTGTTTGC | GATCGCCGCC | GGCTATGCGG | 2050 |
| CCTGGCTGCA | CTTCGCGCCC | CGCGGCGCCG | GCGAAGGGCG | GCTGATCCGC | 2100 |
| GCGCTGACGA | CAGCCCCGGT | ACCGATCGTG | GCCGGTTTCA | TGGCGGCGGT | 2150 |
| GTTTCGTGCG | TCCATGGTGG | CCGGGATCGT | GCGACAGTAC | CCGACCTACT | 2200 |
| CCAACGGCTG | GTCCAACGTG | CGGGCGTTTG | TCGGCGGCTG | CGGACTGGCC | 2250 |
| GACGACGTAC | TCGTGAGGCC | TGATACCAAT | GCGGGTTTCA | TGAAGCCGCT | 2300 |
| GGACGGCGAT | TCGGGTTCTT | GGGGCCCTTT | GGGCCCCTG | GGTGGAGTCA | 2350 |
| ACCCGGTCCG | CTTCACGCCC | AACGGCGTAC | CGGAACACAC | GGTGGCCGAG | 2400 |
| GCGATCGTGA | TGAAACCCAA | CCAGCCCGGC | ACCGACTACG | ACTGGGATGC | 2450 |
| GCCGACCAAG | CTGACGAGTC | CTGGCATCAA | TGTTTCTACG | GTGCCGCTGC | 2500 |
| CCTATGGGCT | CGATCCCGCC | CGGGTACCGT | TGGCAGGCAC | CTACACCACC | 2550 |
| GGCGCACAGC | AACAGAGCAC | ACTCGTCTCG | GCGTGGTATC | TCCTGCCTAA | 2600 |
| GCCGGACGAC | GGGCATCCGC | TGGTCGTGGT | GACCGCCGCG | GGCAAGATCG | 2650 |
| CCGGCAACAG | CGTGCTGCAC | GGGTACACCC | CCGGGCAGAC | TGTGGTGTCT | 2700 |
| GAATACGCCA | TGCCGGGACC | CGGAGCGCTG | GTACCCGCGG | GGCGGATGGT | 2750 |
| GCCCCGACGAC | CTATACGGAG | AGCAGCCCAA | GGCGTGGCGC | AACCTGCGCT | 2800 |
| TCGCCCCGAGC | AAAGATGCCC | GCCGATGCCG | TCGCGGTCCG | GGTGGTGGCC | 2850 |
| GAGGATCTGT | CGCTGACACC | GGAGGACTGG | ATCGCGGTGA | CCCCGCCGCG | 2900 |
| GGTACCGGAC | CTGCGCTCAC | TGCAGGAATA | TGTGGGCTCG | ACGCAGCCGG | 2950 |
| TGCTGCTGGA | CTGGGCGGTC | GGTTTGGCCT | TCCCGTGCCA | GCAGCCGATG | 3000 |
| CTGCACGCCA | ATGGCATCGC | CGAAATCCCG | AAGTTCCGCA | TCACACCGGA | 3050 |
| CTACTCGGCT | AAGAAGCTGG | ACACCGACAC | GTGGGAAGAC | GGCACTAACG | 3100 |
| GCGGCCTGCT | CGGGATCACC | GACCTGTTGC | TGCGGGCCCA | CGTCATGGCC | 3150 |
| ACCTACCTGT | CCCGCGACTG | GGCCCGCGAT | TGGGGTTCCC | TGCGCAAGTT | 3200 |
| CGACACCCTG | GTGATGCCCC | CTCCCGCCCA | GCTCGAGTTG | GGCACC CGA | 3250 |
| CCCGCAGCGG | CCTGTGGTCA | CCGGGCAAGA | TCCGAATTGG | TCCATAG | 3297 |

2) INFORMATION FOR SEQ ID NO: 2106

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2106

GCGAGCAGAG CACGCCCTCC TCGCCGCTCG C

31

2) INFORMATION FOR SEQ ID NO: 2107

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2107

GCGAGCTCCC CATCTCTGGT TGGCAGCTC GC

32

2) INFORMATION FOR SEQ ID NO: 2108

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2108

GCGGGCAACT TCRTCAAGAA GGTTGGTTAC AACCCGCCCG C

41

2) INFORMATION FOR SEQ ID NO: 2109

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

1088

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2109

GCGGGCCCAA TCTCTGGTTG GAAYGGTGAC AAGCCCGC

38

2) INFORMATION FOR SEQ ID NO: 2110

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2110

GCGGGCCCTT AACGATTTCA GCGAATCTGG ATTCAGCCCG C

41

2) INFORMATION FOR SEQ ID NO: 2111

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2111

GCGGGCATGA TTGAAGCCAC CACCAACGCT TCCTGGCCCG C

41

2) INFORMATION FOR SEQ ID NO: 2112

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2112

GCGGGCTTGA TGAAGTTTTG GGTTCCTTG ACAATTGCC CGC

43

2) INFORMATION FOR SEQ ID NO: 2113

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2113

GCGGGCACAA GGGTTGGACT AAGGAAACCA AGGCAGCCCCG C

41

2) INFORMATION FOR SEQ ID NO: 2114

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2114

GCGGGCATCG ATGCTATTGA ACCACCTGTC AGACCGCCCCG C

41

2) INFORMATION FOR SEQ ID NO: 2115

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2115

GCGGGCTTGA TGATTCCTC GAATCTAGAT TGGGCCCCG

39

2) INFORMATION FOR SEQ ID NO: 2116

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2116

GCGGGCGGTA AGTCCACCGG TAAGACCTTG TTGGCCCCG

39

1090

2) INFORMATION FOR SEQ ID NO: 2117

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2117

GCGGGCGACG CCATTGAGCC ACCTTCGAGA GCCCCG

36

2) INFORMATION FOR SEQ ID NO: 2118

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2118

GCGGGCTCCT TGACAATTTC TTCGTATCTG TTCTTGCCCC GC

42

2) INFORMATION FOR SEQ ID NO: 2119

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2119

GCGGGCTTAC AACCCTAAGG CTGTTCCATT CGTTGCCCGC

40

2) INFORMATION FOR SEQ ID NO: 2120

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1091

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2120

GCGGGCAGAA ACCAAGGCTG GTAAGGTTAC CGGTAGCCCCG C 41

2) INFORMATION FOR SEQ ID NO: 2121

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2121

CGACCGCIAG CCGCACACCA AGTTCCGGTC G 31

2) INFORMATION FOR SEQ ID NO: 2122

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2122

CCGAGCGAAT GTAGGAGTCC AGGGTCTCTG CTCGG 35

2) INFORMATION FOR SEQ ID NO: 2123

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2123

TCTACCACTG AAGCATTAC 19

2) INFORMATION FOR SEQ ID NO: 2124

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2124

TAGGTACTGT AGGTTTATTG

20

2) INFORMATION FOR SEQ ID NO: 2125

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2125

CACGCGGATT TTGAATCTCT TCCTCTAGTA GCGCGTG

37

2) INFORMATION FOR SEQ ID NO: 2126

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2126

ATATCAGAGA CTGATGAG

18

2) INFORMATION FOR SEQ ID NO: 2127

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1093

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2127

TAGCATATTC AGAGAATATT GT

22

2) INFORMATION FOR SEQ ID NO: 2128

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2128

CAACGCTGGA GAATCTATAT TTGTAGAAAC TGC GTTG

37

2) INFORMATION FOR SEQ ID NO: 2129

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8133 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium difficile*
- (B) STRAIN: VPI 10463
- (C) ACCESSION NUMBER: X51797

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2129

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGTCTTTAA | TATCTAAAGA | AGAGTTAATA | AAACTCGCAT | ATAGCATTAG | 50 |
| ACCAAGAGAA | AATGAGTATA | AACTATACT | AACTAATTTA | GACGAATATA | 100 |
| ATAAGTTAAC | TACAAACAAT | AATGAAAATA | AATATTTGCA | ATTAAAAAAA | 150 |
| CTAAATGAAT | CAATTGATGT | TTTTATGAAT | AAATATAAAA | CTTCAAGCAG | 200 |
| AAATAGAGCA | CTCTCTAATC | TAAAAAAGA | TATATTAAAA | GAAGTAATTC | 250 |
| TTATTAAAAA | TTCCAATACA | AGCCCTGTAG | AAAAAAATTT | ACATTTTGTA | 300 |
| TGGATAGGTG | GAGAAGTCAG | TGATATTGCT | CTTGAATACA | TAAAACAATG | 350 |
| GGCTGATATT | AATGCAGAAT | ATAATATTAA | ACTGTGGTAT | GATAGTGAAG | 400 |
| CATTCTTAGT | AAATACACTA | AAAAAGGCTA | TAGTTGAATC | TTCTACCACT | 450 |
| GAAGCATTAC | AGCTACTAGA | GGAAGAGATT | CAAAATCCTC | AATTTGATAA | 500 |
| TATGAAATTT | TACAAAAAAA | GGATGGAATT | TATATATGAT | AGACAAAAAA | 550 |
| GGTTTATAAA | TTATTATAAA | TCTCAAATCA | ATAAACCTAC | AGTACCTACA | 600 |
| ATAGATGATA | TTATAAAGTC | TCATCTAGTA | TCTGAATATA | ATAGAGATGA | 650 |
| AACTGTATTA | GAATCATATA | GAACAAATTC | TTTGAGAAAA | ATAAATAGTA | 700 |
| ATCATGGGAT | AGATATCAGG | GCTAATAGTT | TGTTTACAGA | ACAAGAGTTA | 750 |
| TTAAATATTT | ATAGTCAGGA | GTTGTAAAT | CGTGGAATTT | TAGCTGCAGC | 800 |

1094

| | | | | | |
|-------------|-------------|-------------|-------------|------------|------|
| ATCTGACATA | GTAAGATTAT | TAGCCCTAAA | AAATTTTGGC | GGAGTATATT | 850 |
| TAGATGTTGA | TATGCTTCCA | GGTATTCAC | CTGATTTATT | TAAAACAATA | 900 |
| TCTAGACCTA | GCTCTATTGG | ACTAGACCGT | TGGGAAATGA | TAAAATTAGA | 950 |
| GGCTATTATG | AAGTATAAAA | AATATATAAA | TAATTATACA | TCAGAAAAC | 1000 |
| TTGATAAACT | TGATCAACAA | TTAAAAGATA | ATTTTAAACT | CATTATAGAA | 1050 |
| AGTAAAAGTG | AAAAATCTGA | GATATTTTCT | AAATTAGAAA | ATTTAAATGT | 1100 |
| ATCTGATCTT | GAAATTAAAA | TAGCTTTCGC | TTTAGGCAGT | GTTATAAATC | 1150 |
| AAGCCTTGAT | ATCAAAACAA | GGTTCATATC | TTACTAACCT | AGTAATAGAA | 1200 |
| CAAGTAAAAA | ATAGATATCA | ATTTTTTAAAC | CAACACCTTA | ACCCAGCCAT | 1250 |
| AGAGTCTGAT | AATAACTTCA | CAGATACTAC | TAAAATTTTT | CATGATTCAT | 1300 |
| TATTTAATTC | AGCTACCGCA | GAAAACTCTA | TGTTTTTAAAC | AAAAATAGCA | 1350 |
| CCATACTTAC | AAGTAGGTTT | TATGCCAGAA | GCTCGCTCCA | CAATAAGTTT | 1400 |
| AAGTGGTCCA | GGAGCTTATG | CGTCAGCTTA | CTATGATTTT | ATAAATTTAC | 1450 |
| AAGAAAATAC | TATAGAAAAA | ACTTTTAAAG | CATCAGATTT | AATAGAATTT | 1500 |
| AAATTCCCAG | AAAATAATCT | ATCTCAATTG | ACAGAACAAG | AAATAAATAG | 1550 |
| TCTATGGAGC | TTTGATCAAG | CAAGTGCAAA | ATATCAATTT | GAGAAATATG | 1600 |
| TAAGAGATTA | TACTGGTGGG | TCTCTTTCTG | AAGACAATGG | GGTAGACTTT | 1650 |
| AATAAAAAATA | CTGCCCTCGA | CAAAAACTAT | TTATTAAATA | ATAAAATTCC | 1700 |
| ATCAACAAT | GTAGAAGAAG | CTGGAAGTAA | AAATTATGTT | CATTATATCA | 1750 |
| TACAGTTACA | AGGAGATGAT | ATAAGTTATG | AAGCAACATG | CAATTTATTT | 1800 |
| TCTAAAAATC | CTAAAAATAG | TATTATTATA | CAACGAAATA | TGAATGAAAG | 1850 |
| TGCAAAAAGC | TACTTTTTTAA | GTGATGATGG | AGAATCTATT | TTAGAATTAA | 1900 |
| ATAAATATAG | GATACCTGAA | AGATTAAAAA | ATAAGGAAAA | AGTAAAAGTA | 1950 |
| ACCTTTTATTG | GACATGGTAA | AGATGAATTC | AACACAAGCG | AATTTGCTAG | 2000 |
| ATTAAGTGTA | GATTCACTTT | CCAATGAGAT | AAGTTCATTT | TTAGATACCA | 2050 |
| TAAAATTAGA | TATATCACCT | AAAAATGTAG | AAGTAAACTT | ACTTGGATGT | 2100 |
| AATATGTTTA | GTTATGATTT | TAATGTTGAA | GAAACTTATC | CTGGGAAGTT | 2150 |
| GCTATTAAGT | ATTATGGACA | AAATTACTTC | CACCTTACCT | GATGTAAATA | 2200 |
| AAAATTCTAT | TACTATAGGA | GCAAATCAAT | ATGAAGTAAG | AATTAATAGT | 2250 |
| GAGGGAAGAA | AAGAACTTCT | GGCTCACTCA | GGTAAATGGA | TAAATAAAGA | 2300 |
| AGAAGCTATT | ATGAGCGATT | TATCTAGTAA | AGAATACATT | TTTTTTGATT | 2350 |
| CTATAGATAA | TAAGCTAAAA | GCAAAGTCCA | AGAATATTCC | AGGATTAGCA | 2400 |
| TCAATATCAG | AAGATATAAA | AACATTATTA | CTTGATGCAA | GTGTTAGTCC | 2450 |
| TGATACAAAA | TTTATTTTAA | ATAATCTTAA | GCTTAATATT | GAATCTTCTA | 2500 |
| TTGGGGATTA | CATTTATTAT | GAAAAATTAG | AGCCTGTTAA | AAATATAATT | 2550 |
| CACAATTCTA | TAGATGATTT | AATAGATGAG | TTCAATCTAC | TTGAAAATGT | 2600 |
| ATCTGATGAA | TTATATGAAT | TAAAAAAATT | AAATAATCTA | GATGAGAAGT | 2650 |
| ATTTAATATC | TTTTGAAGAT | ATCTCAAAAA | ATAATTCAAC | TTACTCTGTA | 2700 |
| AGATTTATTA | ACAAAAGTAA | TGGTGAGTCA | GTTTATGTAG | AAACAGAAAA | 2750 |
| AGAAATTTTT | TCAAAATATA | GCGAACATAT | TACAAAAGAA | ATAAGTACTA | 2800 |
| TAAAGAATAG | TATAATTACA | GATGTTAATG | GTAATTTATT | GGATAATATA | 2850 |
| CAGTTAGATC | ATACTTCTCA | AGTTAATACA | TTAAACGCAG | CATTCTTTAT | 2900 |
| TCAATCATTA | ATAGATTATA | GTAGCAATAA | AGATGTACTG | AATGATTTAA | 2950 |
| GTACCTCAGT | TAAGGTTCAA | CTTTATGCTC | AACATTTTAG | TACAGGTTTA | 3000 |
| AATACTATAT | ATGACTCTAT | CCAATTAGTA | AATTTAATAT | CAAATGCAGT | 3050 |
| AAATGATACT | ATAAATGTAC | TACCTACAAT | AACAGAGGGG | ATACCTATTG | 3100 |
| TATCTACTAT | ATTAGACGGA | ATAAACTTAG | GTGCAGCAAT | TAAGGAATTA | 3150 |
| CTAGACGAAC | ATGACCCATT | ACTAAAAAAA | GAATTAGAAG | CTAAGGTGGG | 3200 |
| TGTTTTAGCA | ATAAATATGT | CATTATCTAT | AGCTGCAACT | GTAGCTTCAA | 3250 |
| TTGTTGGAAT | AGGTGCTGAA | GTTACTATTT | TCTTATTACC | TATAGCTGGT | 3300 |
| ATATCTGCAG | GAATACCTTC | ATTAGTTAAT | AATGAATTAA | TATTGCATGA | 3350 |
| TAAGGCAACT | TCAGTGGTAA | ACTATTTTAA | TCATTTGTCT | GAATCTAAAA | 3400 |
| AATATGGCCC | TCTTAAAACA | GAAGATGATA | AAATTTTAGT | TCCTATTGAT | 3450 |
| GATTTAGTAA | TATCAGAAAT | AGATTTTAAT | AATAATTCGA | TAAAAC | 3500 |
| AACATGTAAT | ATATTAGCAA | TGGAGGGGGG | ATCAGGACAC | ACAGTGACTG | 3550 |
| GTAATATAGA | TCACCTTTTC | TCATCTCCAT | CTATAAGTTC | TCATATTCCT | 3600 |
| TCATTATCAA | TTTATTCTGC | AATAGGTATA | GAAACAGAAA | ATCTAGATTT | 3650 |

| | | | | | |
|------------|------------|-------------|------------|------------|------|
| TTCAAAAAA | ATAATGATGT | TACCTAATGC | TCCTTCAAGA | GTGTTTTGGT | 3700 |
| GGGAAACTGG | AGCAGTTCCA | GGTTTAAGAT | CATTGGAAAA | TGACGGAACT | 3750 |
| AGATTACTTG | ATTCAATAAG | AGATTTTATAC | CCAGGTAAAT | TTTACTGGAG | 3800 |
| ATTCTATGCT | TTTTTCGATT | ATGCAATAAC | TACATTAAAA | CCAGTTTATG | 3850 |
| AAGACACTAA | TATTAAAATT | AAACTAGATA | AAGATACTAG | AAACTTCATA | 3900 |
| ATGCCAACTA | TAACACTAA | CGAAATTAGA | AACAAATTAT | CTTATTCATT | 3950 |
| TGATGGAGCA | GGAGGAACTT | ACTCTTTATT | ATTATCTTCA | TATCCAATAT | 4000 |
| CAACGAATAT | AAATTTATCT | AAAGATGATT | TATGGATATT | TAATATTGAT | 4050 |
| AATGAAGTAA | GAGAAATATC | TATAGAAAAT | GGTACTATTA | AAAAAGGAAA | 4100 |
| GTTAATAAAA | GATGTTTTAA | GTAAAATTGA | TATAAATAAA | AATAAACTTA | 4150 |
| TTATAGGCAA | TCAAACAATA | GATTTTTTCAG | GCGATATAGA | TAATAAAGAT | 4200 |
| AGATATATAT | TCTTGACTTG | TGAGTTAGAT | GATAAAATTA | GTTTAATAAT | 4250 |
| AGAAATAAAT | CTTGTTGCAA | AATCTTATAG | TTTGTTATTG | TCTGGGGATA | 4300 |
| AAAATTATTT | GATATCCAAT | TTATCTAATA | CTATTGAGAA | AATCAATACT | 4350 |
| TTAGGCCTAG | ATAGTAAAAA | TATAGCGTAC | AATTACACTG | ATGAATCTAA | 4400 |
| TAATAAATAT | TTTGGAGCTA | TATCTAAAAC | AAGTCAAAAA | AGCATAATAC | 4450 |
| ATTATAAAAA | AGACAGTAAA | AATATATTAG | AATTTTATAA | TGACAGTACA | 4500 |
| TTAGAATTTA | ACAGTAAAGA | TTTTATTGCT | GAAGATATAA | ATGTATTTAT | 4550 |
| GAAAGATGAT | ATTAATACTA | TAACAGGAAA | ATACTATGTT | GATAATAATA | 5000 |
| CTGATAAAAG | TATAGATTTT | TCTATTTCTT | TAGTTAGTAA | AAATCAAGTA | 5050 |
| AAAGTAAATG | GATTATATTT | AAATGAATCC | GTATACTCAT | CTTACCTTGA | 5100 |
| TTTTGTGAAA | AATTCAGATG | GACACCATAA | TACTTCTAAT | TTTATGAATT | 5150 |
| TATTTTTTGA | CAATATAAGT | TTCTGGAAAT | TGTTTGGGTT | TGAAAATATA | 5200 |
| AATTTTGTAA | TCGATAAATA | CTTTACCCTT | GTTGGTAAAA | CTAATCTTGG | 5250 |
| ATATGTAGAA | TTTATTTGTG | ACAATAATAA | AAATATAGAT | ATATATTTTG | 5300 |
| GTGAATGGAA | AACATCGTCA | TCTAAAAGCA | CTATATTTAG | CGGAAATGGT | 5350 |
| AGAAATGTTG | TAGTAGAGCC | TATATATAAT | CCTGATACGG | GTGAAGATAT | 5400 |
| ATCTACTTCA | CTAGATTTTT | CCTATGAACC | TCTCTATGGA | ATAGATAGAT | 5450 |
| ATATAAATAA | AGTATTGATA | GCACCTGATT | TATATACAAG | TTTAATAAAT | 5500 |
| ATTAATACCA | ATTATTATTC | AAATGAGTAC | TACCCTGAGA | TTATAGTTCT | 5550 |
| TAACCCAAAT | ACATTCCACA | AAAAAGTAAA | TATAAATTTA | GATAGTTCTT | 5600 |
| CTTTTGAGTA | TAAATGGTCT | ACAGAAGGAA | GTGACTTTAT | TTTAGTTAGA | 5650 |
| TACTTAGAAG | AAAGTAATAA | AAAAATATTA | CAAAAAATAA | GAATCAAAGG | 5700 |
| TATCTTATCT | AATACTCAAT | CATTTAATAA | AATGAGTATA | GATTTTAAAG | 5750 |
| ATATTAAAAA | ACTATCATT | GGATATATAA | TGAGTAATTT | TAAATCATT | 5800 |
| AATTCTGAAA | ATGAATTAGA | TAGAGATCAT | TTAGGATTTA | AAATAATAGA | 5850 |
| TAATAAAACT | TATTACTATG | ATGAAGATAG | TAAATTAGTT | AAAGGATTAA | 5900 |
| GCAAGTAATT | GACAATAATA | AGTATTATTT | CAATCCTGAC | ACTGCTATCA | 5950 |
| TCTCAAAAGG | TTGGCAGACT | GTTAATGGTA | GTAGATACTA | CTTTGATACT | 6000 |
| GATACCGCTA | TTGCCTTTAA | TGGTTATAAA | ACTATTGATG | GTAAACACTT | 6050 |
| TTATTTTGAT | AGTGATTGTG | TAGTGAAAAT | AGGTGTGTTT | AGTACCTCTA | 6100 |
| ATGGATTTGA | ATATTTTGCA | CCTGCTAATA | CTTATAATAA | TAACATAGAA | 6150 |
| GGTCAGGCTA | TAGTTTATCA | AAGTAAATTC | TTAACTTTGA | ATGGTAAAAA | 6200 |
| ATATTACTTT | GATAATAACT | CAAAAGCAGT | TACCGGATTG | CAAACATTG | 6250 |
| ATAGTAAAAA | ATATTACTTT | AATACTAACA | CTGCTGAAGC | AGCTACTGGA | 6300 |
| TGGCAAACTA | TTGATGGTAA | AAAATATTAC | TTTAATACTA | ACACTGCTGA | 6350 |
| AGCAGCTACT | GGATGGCAAA | CTATTGATGG | TAAAAAATAT | TACTTTAATA | 6400 |
| CTAACACTGC | TATAGCTTCA | ACTGGTTATA | CAATTATTAA | TGGTAAACAT | 6450 |
| TTTTATTTTA | ATACTGATGG | TATTATGCAG | ATAGGAGTGT | TTAAAGGACC | 6500 |
| TAATGGATTT | GAATATTTTG | CACCTGCTAA | TACGGATGCT | AACAACATAG | 6550 |
| AAGGTCAAGC | TATACTTTAC | CAAAATGAAT | TCTTAACTTT | GAATGGTAAA | 6600 |
| AAATATTACT | TTGGTAGTGA | CTCAAAAGCA | GTTACTGGAT | GGAGAATTAT | 6650 |
| TAACAATAAG | AAATATTACT | TTAATCCTAA | TAATGCTATT | GCTGCAATTC | 6700 |
| ATCTATGCAC | TATAAATAAT | GACAAGTATT | ACTTTAGTTA | TGATGGAATT | 6750 |
| CTTCAAAATG | GATATATTAC | TATTGAAAGA | AATAATTTCT | ATTTTGATGC | 6800 |
| TAATAATGAA | TCTAAAATGG | TAACAGGAGT | ATTAAAGGA | CCTAATGGAT | 6850 |
| TTGAGTATTT | TGCACCTGCT | AATACTCACA | ATAATAACAT | AGAAGGTCAG | 6900 |

| | | | | | |
|------------|------------|-------------|-------------|-------------|------|
| GCTATAGTTT | ACCAGAACAA | ATTCTTAACT | TTGAATGGCA | AAAAATATTA | 6950 |
| TTTTGATAAT | GACTCAAAAG | CAGTTACTGG | ATGGCAAACC | ATTGATGGTA | 7000 |
| AAAAATATTA | CTTTAATCTT | AACACTGCTG | AAGCAGCTAC | TGGATGGCAA | 7050 |
| ACTATTGATG | GTAAAAAATA | TTACTTTAAT | CTTAACACTG | CTGAAGCAGC | 7100 |
| TACTGGATGG | CAAACTATTG | ATGGTAAAAA | ATATTACTTT | AATACTAACA | 7150 |
| CTTTCATAGC | CTCAACTGGT | TATACAAGTA | TTAATGGTAA | ACATTTTTTAT | 7200 |
| TTTAATACTG | ATGGTATTAT | GCAGATAGGA | GTGTTTAAAG | GACCTAATGG | 7250 |
| ATTTGAATAC | TTTGACCTG | CTAATACGGA | TGCTAACAAAC | ATAGAAGGTC | 7300 |
| AAGCTATACT | TTACCAAAAT | AAATTCTTAA | CTTTGAATGG | TAAAAAATAT | 7350 |
| TACTTTGGTA | GTGACTCAA | AGCAGTTACC | GGACTGCGAA | CTATTGATGG | 7400 |
| TAAAAAATAT | TACTTTAATA | CTAACACTGC | TGTTGCAGTT | ACTGGATGGC | 7450 |
| AAACTATTAA | TGGTAAAAAA | TACTACTTTA | ATACTAACAC | TTCTATAGCT | 7500 |
| TCAACTGGTT | ATACAATTAT | TAGTGGTAAA | CATTTTTTATT | TTAATACTGA | 7550 |
| TGGTATTATG | CAGATAGGAG | TGTTTAAAGG | ACCTGATGGA | TTTGAATACT | 7600 |
| TTGCACCTGC | TAATACAGAT | GCTAACAAATA | TAGAAGGTCA | AGCTATACGT | 7650 |
| TATCAAAATA | GATTCCTATA | TTTACATGAC | AATATATATT | ATTTTGGTAA | 7700 |
| TAATTCAAAA | GCGGCTACTG | GTTGGGTAAC | TATTGATGGT | AATAGATATT | 7750 |
| ACTTCGAGCC | TAATACAGCT | ATGGGTGCGA | ATGGTTATAA | AACTATTGAT | 7800 |
| AATAAAAAAT | TTTACTTTAG | AAATGGTTTA | CCTCAGATAG | GAGTGTTTAA | 7850 |
| AGGGTCTAAT | GGATTTGAAT | ACTTTGCACC | TGCTAATACG | GATGCTAACA | 7900 |
| ATATAGAAGG | TCAAGCTATA | CGTTATCAAA | ATAGATTCCT | ACATTTACTT | 7950 |
| GGAAAAATAT | ATTACTTTGG | TAATAATTCA | AAAGCAGTTA | CTGGATGGCA | 8000 |
| AACTATTAAT | GGTAAAGTAT | ATTACTTTAT | GCCTGATACT | GCTATGGCTG | 8050 |
| CAGCTGGTGG | ACTTTTCGAG | ATTGATGGTG | TTATATATTT | CTTTGGTGTT | 8100 |
| GATGGAGTAA | AAGCCCCTGG | GATATATGGC | TAA | | 8133 |

2) INFORMATION FOR SEQ ID NO: 2130

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7101 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium difficile*
- (B) STRAIN: VPI 10463
- (C) ACCESSION NUMBER: X53138

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2130

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGAGTTTAG | TTAATAGAAA | ACAGTTAGAA | AAAATGGCAA | ATGTAAGATT | 50 |
| TCGTACTCAA | GAAGATGAAT | ATGTTGCAAT | ATTGGATGCT | TTAGAAGAAT | 100 |
| ATCATAATAT | GTCAGAGAAT | ACTGTAGTCG | AAAAATATTT | AAAATTAAAA | 150 |
| GATATAAATA | GTTTAACAGA | TATTTATATA | GATACATATA | AAAAATCTGG | 200 |
| TAGAAATAAA | GCCTTAAAAA | AATTTAAGGA | ATATCTAGTT | ACAGAAGTAT | 250 |
| TAGAGCTAAA | GAATAATAAT | TTAACTCCAG | TTGAGAAAAA | TTTACATTTT | 300 |
| GTTTGGATTG | GAGGTCAAAT | AAATGACACT | GCTATTAATT | ATATAAATCA | 350 |
| ATGGAAAGAT | GTAAATAGTG | ATTATAATGT | TAATGTTTTT | TATGATAGTA | 400 |
| ATGCATTTT | GATAAACACA | TTGAAAAAAA | CTGTAGTAGA | ATCAGCAATA | 450 |
| AATGATACAC | TTGAATCATT | TAGAGAAAAC | TTAAATGACC | CTAGATTTGA | 500 |
| CTATAATAAA | TTCTTCAGAA | AACGTATGGA | AATAATTTAT | GATAAACAGA | 550 |
| AAAATTTT | CACTAT | AAAGCTCAA | GAGAAGAAAA | TCCTGAACCT | 600 |

| | | | | | |
|------------|-------------|------------|------------|-------------|------|
| ATAATTGATG | ATATTGTAAA | GACATATCTT | TCAAATGAGT | ATTCAAAGGA | 650 |
| GATAGATGAA | CTTAATACCT | ATATTGAAGA | ATCCTTAAAT | AAAAATTACAC | 700 |
| AGAATAGTGG | AAATGATGTT | AGAAACTTTG | AAGAATTTAA | AAATGGAGAG | 750 |
| TCATTCAACT | TATATGAACA | AGAGTTGGTA | GAAAGGTGGA | ATTTAGCTGC | 800 |
| TGCTTCTGAC | ATATTAAGAA | TATCTGCATT | AAAAGAAATT | GGTGGTATGT | 850 |
| ATTTAGATGT | TGATATGTTA | CCAGGAATAC | AACCAGACTT | ATTTGAGTCT | 900 |
| ATAGAGAAAC | CTAGTTCAGT | AACAGTGGAT | TTTTGGGAAA | TGACAAAGTT | 950 |
| AGAAGCTATA | ATGAAATACA | AAGAATATAT | ACCAGAATAT | ACCTCAGAAC | 1000 |
| ATTTTGACAT | GTTAGACGAA | GAAGTTCAAA | GTAGTTTTGA | ATCTGTTCTA | 1050 |
| GCTTCTAAGT | CAGATAAATC | AGAAATATTC | TCATCACTTG | GTGATATGGA | 1100 |
| GGCATCACCA | CTAGAAAGTTA | AAATTGCATT | TAATAGTAAG | GGTATTATAA | 1150 |
| ATCAAGGGCT | AATTTCTGTG | AAAGACTCAT | ATTGTAGCAA | TTTAATAGTA | 1200 |
| AAACAAATCG | AGAATAGATA | TAAAATATTG | AATAATAGTT | TAAATCCAGC | 1250 |
| TATTAGCGAG | GATAATGATT | TTAATACTAC | AACGAATACC | TTTATTGATA | 1300 |
| GTATAATGGC | TGAAGCTAAT | GCAGATAATG | GTAGATTTAT | GATGGAACTA | 1350 |
| GGAAAGTATT | TAAGAGTTGG | TTTCTTCCCA | GATGTTAAAA | CTACTATTAA | 1400 |
| CTTAAGTGGC | CCTGAAGCAT | ATGCGGCAGC | TTATCAAGAT | TTATTAATGT | 1450 |
| TTAAAGAAGG | CAGTATGAAT | ATCCATTTGA | TAGAAGCTGA | TTTAAGAAAC | 1500 |
| TTTGAAATCT | CTAAAACTAA | TATTTCTCAA | TCAACTGAAC | AAGAAATGGC | 1550 |
| TAGCTTATGG | TCATTTGACG | ATGCAAGAGC | TAAAGCTCAA | TTTGAAGAAT | 1600 |
| ATAAAAGGAA | TTATTTTGAA | GGTTCCTCTG | GTGAAGATGA | TAATCTTGAT | 1650 |
| TTTTCTCAAA | ATATAGTAGT | TGACAAGGAG | TATCTTTTAG | AAAAAATATC | 1700 |
| TTCATTAGCA | AGAAGTTCAG | AGAGAGGATA | TATACACTAT | ATTGTTTCAGT | 1750 |
| TACAAGGAGA | TAAAATTAGT | TATGAAGCAG | CATGTAACCT | ATTTGCAAAG | 1800 |
| ACTCCTTATG | ATAGTGTACT | GTTTCAGAAA | AATATAGAAG | ATTCAGAAAT | 1850 |
| TGCATATTAT | TATAATCCTG | GAGATGGTGA | AATACAAGAA | ATAGACAAGT | 1900 |
| ATAAAATTCC | AAGTATAATT | TCTGATAGAC | CTAAGATTAA | ATTAACATTT | 1950 |
| ATTGGTCATG | GTAAAGATGA | ATTTAATACT | GATATATTTG | CAGGTTTTGA | 2000 |
| TGTAGATTCA | TTATCCACAG | AAATAGAAGC | AGCAATAGAT | TTAGCTAAAG | 2050 |
| AGGATATTTT | TCCTAAGTCA | ATAGAAATAA | ATTTATTAGG | ATGTAATATG | 2100 |
| TTTAGCTACT | CTATCAACGT | AGAGGAGACT | TATCCTGGAA | AATTATTACT | 2150 |
| TAAAGTTAAA | GATAAAATAT | CAGAATTAAT | GCCATCTATA | AGTCAAGACT | 2200 |
| CTATTATAGT | AAGTGCAAAT | CAATATGAAG | TTAGAATAAA | TAGTGAAGGA | 2250 |
| AGAAGAGAAT | TATTGGATCA | TTCTGGTGAA | TGGATAAATA | AAGAAGAAAG | 2300 |
| TATTATAAAG | GATATTTTCAT | CAAAAGAATA | TATATCATTT | AATCCTAAAG | 2350 |
| AAAATAAAAT | TACAGTAAAA | TCTAAAAATT | TACCTGAGCT | ATCTACATTA | 2400 |
| TTACAAGAAA | TTAGAAATAA | TTCTAATTCA | AGTGATATTG | AACTAGAAGA | 2450 |
| AAAAGTAATG | TTAACAGAAT | GTGAGATAAA | TGTTATTTCA | AATATAGATA | 2500 |
| CGCAAATTGT | TGAGGAAAGG | ATTGAAGAAG | CTAAGAATTT | AACTTCTGAC | 2550 |
| TCTATTAATT | ATATAAAAGA | TGAATTTAAA | CTAATAGAAT | CTATTTCTGA | 2600 |
| TGCACTATGT | GACTTAAAAC | AACAGAATGA | ATTAGAAGAT | TCTCATTTTA | 2650 |
| TATCTTTTGA | GGACATATCA | GAGACTGATG | AGGGATTTAG | TATAAGATTT | 2700 |
| ATTAATAAAG | AAACTGGAGA | ATCTATATTT | GTAGAACTG | AAAAACAAT | 2750 |
| ATTCTCTGAA | TATGCTAATC | ATATAACTGA | AGAGATTTCT | AAGATAAAAG | 2800 |
| GTAATAATAT | TGATACTGTA | AATGGTAAGT | TAGTAAAAAA | AGTAAATTTA | 2850 |
| GATACTACAC | ACGAAGTAAA | TACTTTAAAT | GCTGCATTTT | TTATACAATC | 2900 |
| ATTAATAGAA | TATAATAGTT | CTAAAGAATC | TCTTAGTAAT | TTAAGTGTAG | 2950 |
| CAATGAAAGT | CCAAGTTTAC | GCTCAATTAT | TTAGTACTGG | TTTAAATACT | 3000 |
| ATTACAGATG | CAGCCAAAGT | TGTTGAATTA | GTATCAACTG | CATTAGATGA | 3050 |
| AACTATAGAC | TTACTTCCTA | CATTATCTGA | AGGATTACCT | ATAATTGCAA | 3100 |
| CTATTATAGA | TGGTGTAAGT | TTAGGTGCAG | CAATCAAAGA | GCTAAGTGAA | 3150 |
| ACGAGTGACC | CATTATTAAG | ACAAGAAATA | GAAGCTAAGA | TAGGTATAAT | 3200 |
| GGCAGTAAAT | TTAACAACAG | CTACAACCTG | AATCATTACT | TCATCTTTGG | 3250 |
| GGATAGCTAG | TGGATTTAGT | ATACTTTTAG | TTCTTTTAGC | AGGAATTTCA | 3300 |
| GCAGGTATAC | CAAGCTTAGT | AAACAATGAA | CTTGTAATTC | GAGATAAGGC | 3350 |
| AACAAAGGTT | GTAGATTATT | TTAAACATGT | TTTATTAGTT | GAAACTGAAG | 3400 |
| GAGTATTTAC | TTTATTAGAT | GATAAAATAA | TGATGCCACA | AGATGATTTA | 3450 |

| | | | | | |
|-------------|-------------|------------|-------------|-------------|------|
| GTGATATCAG | AAATAGATTT | TAATAATAAT | TCAATAGTTT | TAGGTAAATG | 3500 |
| TGAAATCTGG | AGAATGGAAG | GTGGTTCAGG | TCATACTGTA | ACTGATGATA | 3550 |
| TAGATCACTT | CTTTTCAGCA | CCATCAATAA | CATATAGAGA | GCCACACTTA | 3600 |
| TCTATATATG | ACGTATTGGA | AGTACAAAAA | GAAGAACTTG | ATTTGTCAAA | 3650 |
| AGATTTAATG | GTATTACCTA | ATGCTCCAAA | TAGAGTATTT | GCTTGGGAAA | 3700 |
| CAGGATGGAC | ACCAGGTTTA | AGAAGCTTAG | AAAATGATGG | CACAAAACCTG | 3750 |
| TTAGACCGTA | TAAGAGATAA | CTATGAAGGT | GAGTTTTATT | GGAGATATTT | 3800 |
| TGCTTTTATA | GCTGATGCTT | TAATAACAAC | ATTAATAACCA | AGATATGAAG | 3850 |
| ATACTAATAT | AAGAATAAAT | TTAGATAGTA | ATACTAGAAG | TTTTATAGTT | 3900 |
| CCAATAATAA | CTACAGAATA | TATAAGAGAA | AAATTATCAT | ATTCTTTCTA | 3950 |
| TGGTTCAGGA | GGAACCTTATG | CATTGTCTCT | TTCTCAATAT | AATATGGGTA | 4000 |
| TAAATATAGA | ATTAAGTGAA | AGTGATGTTT | GGATTATAGA | TGTTGATAAT | 4050 |
| GTTGTGAGAG | ATGTAACCTAT | AGAATCTGAT | AAAATTAAAA | AAGGTGATTT | 4100 |
| AATAGAAGGT | ATTTTATCTA | CACTAAGTAT | TGAAGAGAAT | AAAATTATCT | 4150 |
| TAAATAGCCA | TGAGATTAAT | TTTTCTGGTG | AGGTAAATGG | AAGTAATGGA | 4200 |
| TTTGTTTCTT | TAACATTTTC | AATTTTAGAA | GGAATAAATG | CAATTATAGA | 4250 |
| AGTTGATTTA | TTATCTAAAT | CATATAAATT | ACTTATTTCT | GGCGAATTAA | 4300 |
| AAATATTGAT | GTTAAATTCA | AATCATATTC | AACAGAAAAT | AGATTATATA | 4350 |
| GGATTCAATA | GCGAATTACA | GAAAAATATA | CCATATAGCT | TTGTAGATAG | 4400 |
| TGAAGGAAAA | GAGAATGGTT | TTATTAATGG | TTCAACAAAA | GAAGGTTTAT | 4450 |
| TTGTATCTGA | ATTACCTGAT | GTAGTCTT | TAAGTAAGGT | TTATATGGAT | 4500 |
| GATAGTAAGC | CTTCATTGG | ATATTATAGT | AATAATTTGA | AAGATGTCAA | 4550 |
| AGTTATAACT | AAAGATAATG | TTAATATATT | AACAGGTTAT | TATCTTAAGG | 4600 |
| ATGATATAAA | AATCTCTCTT | TCTTTGACTC | TACAAGATGA | AAAACTATA | 4650 |
| AAGTTAAATA | GTGTGCATTT | AGATGAAAGT | GGAGTAGCTG | AGATTTTGAA | 4700 |
| GTTCATGAAT | AGAAAAGGTA | ATACAAATAC | TTCAGATTCT | TTAATGAGCT | 4750 |
| TTTTAGAAAG | TATGAATATA | AAAAGTATTT | TCGTTAATTT | CTTACAATCT | 4800 |
| AATATTAAGT | TTATATTAGA | TGCTAATTTT | ATAATAAGTG | GTACTACTTC | 4850 |
| TATTGGCCAA | TTTGAGTTTA | TTTGTGATGA | AAATGATAAT | ATACAACCAT | 4900 |
| ATTTTCATTAA | GTTTAATACA | CTAGAAACTA | ATTATACTTT | ATATGTAGGA | 4950 |
| AATAGACAAA | ATATGATAGT | GGAACCAAAT | TATGATTTAG | ATGATTCTGG | 5000 |
| AGATATATCT | TCAACTGTTA | TCAATTTCTC | TCAAAAGTAT | CTTTATGGAA | 5050 |
| TAGACAGTTG | TGTTAATAAA | GTTGTAATTT | CACCAAATAT | TTATACAGAT | 5100 |
| GAAATAAATA | TAACGCCTGT | ATATGAAACA | AATAATACTT | ATCCAGAAGT | 5150 |
| TATTGTATTA | GATGCAAATT | ATATAAATGA | AAAAATAAAT | GTTAATATCA | 5200 |
| ATGATCTATC | TATACGATAT | GTATGGAGTA | ATGATGGTAA | TGATTTTATT | 5250 |
| CTTATGTCAA | CTAGTGAAGA | AAATAAGGTG | TCACAAGTTA | AAATAAGATT | 5300 |
| CGTTAATGTT | TTTAAAGATA | AGACTTTGGC | AAATAAGCTA | TCTTTTAACT | 5350 |
| TTAGTGATAA | ACAAGATGTA | CCTGTAAGTG | AAATAATCTT | ATCATTTACA | 5400 |
| CCTTCATATT | ATGAGGATGG | ATTGATTGGC | TATGATTTGG | GTCTAGTTTC | 5450 |
| TTTATATAAT | GAGAAATTTT | ATATTAATAA | CTTTGGAATG | ATGGTATCTG | 5500 |
| GATTAATATA | TATTAATGAT | TCATTATATT | ATTTTAAACC | ACCAGTAAAT | 5550 |
| AATTTGATAA | CTGGATTGTT | GACTGTAGGC | GATGATAAAT | ACTACTTTAA | 5600 |
| TCCAATTAAT | GGTGGAGCTG | CTTCAATTGG | AGAGACAATA | ATTGATGACA | 5650 |
| AAAATTATTA | TTTCAACCAA | AGTGGAGTGT | TACAAACAGG | TGTATTTAGT | 5700 |
| ACAGAAGATG | GATTTAAATA | TTTTGCCCCA | GCTAATACAC | TTGATGAAAA | 5750 |
| CCTAGAAGGA | GAAGCAATTG | ATTTTACTGG | AAAATTAATT | ATTGACGAAA | 5800 |
| ATATTTATTA | TTTTGATGAT | AATTATAGAG | GAGCTGTAGA | ATGGAAAGAA | 5850 |
| TTAGATGGTG | AAATGCACTA | TTTTAGCCCA | GAAACAGGTA | AAGCTTTTAA | 5900 |
| AGGTCTAAAT | CAAATAGGTG | ATTATAAATA | CTATTTCAAT | TCTGATGGAG | 5950 |
| TTATGCAAAA | AGGATTGT | AGTATAAATG | ATAATAAACA | CTATTTTGAT | 6000 |
| GATTCTGGTG | TTATGAAAGT | AGGTTACACT | GAAATAGATG | GCAAGCATT | 6050 |
| CTACTTTGCT | GAAAACGGAG | AAATGCAAAT | AGGAGTATTT | AATACAGAAG | 6100 |
| ATGGATTTAA | ATATTTTGCT | CATCATAATG | AAGATTTAGG | AAATGAAGAA | 6150 |
| GGTGAAGAAA | TCTCATATTC | TGGTATATTA | AATTTCAATA | ATAAAATTTA | 6200 |
| CTATTTTGAT | GATTCATTTA | CAGCTGTAGT | TGGATGGAAA | GATTTAGAGG | 6250 |
| ATGGTTCAAA | GTATTATTTT | GATGAAGATA | CAGCAGAAGC | ATATATAGGT | 6300 |

| | | | | | |
|------------|-------------|------------|-------------|------------|------|
| TTGTCATTAA | TAAATGATGG | TCAATATTAT | TTTAATGATG | ATGGAATTAT | 6350 |
| GCAAGTTGGA | TTTGTCACCTA | TAAATGATAA | AGTCTTCTAC | TTCTCTGACT | 6400 |
| CTGGAATTAT | AGAATCTGGA | GTACAAAACA | TAGATGACAA | TTATTTCTAT | 6450 |
| ATAGATGATA | ATGGTATAGT | TCAAATTGGT | GTATTTGATA | CTTCAGATGG | 6500 |
| ATATAAATAT | TTTGACACCTG | CTAATACTGT | AAATGATAAT | ATTTACGGAC | 6550 |
| AAGCAGTTGA | ATATAGTGGT | TTAGTTAGAG | TTGGGGAAGA | TGTATATTAT | 6600 |
| TTTGAGAGAA | CATATACAAT | TGAGACTGGA | TGGATATATG | ATATGGAGAA | 6650 |
| TGAAAGTGAT | AAATATTATT | TCAATCCAGA | AACTAAAAAA | GCATGCAAAG | 6700 |
| GTATTAATTT | AATTGATGAT | ATAAAATATT | ATTTTGATGA | GAAGGGCATA | 6750 |
| ATGAGAACGG | GTCTTATATC | ATTTGAAAAT | AATAATTATT | ACTTTAATGA | 6800 |
| GAATGGTGAA | ATGCAATTTG | GTTATATAAA | TATAGAAGAT | AAGATGTTCT | 6850 |
| ATTTTGGTGA | AGATGGTGTC | ATGCAGATTG | GAGTATTTAA | TACACCAGAT | 6900 |
| GGATTTAAAT | ACTTTGCACA | TCAAAATACT | TTGGATGAGA | ATTTTGAGGG | 6950 |
| AGAATCAATA | AACTATACTG | GTTGGTTAGA | TTTAGATGAA | AAGAGATATT | 7000 |
| ATTTTACAGA | TGAATATATT | GCAGCAACTG | GTTTCAGTTAT | TATTGATGGT | 7050 |
| GAGGAGTATT | ATTTTGATCC | TGATACAGCT | CAATTAGTGA | TTAGTGAATA | 7100 |
| G | | | | | 7101 |

2) INFORMATION FOR SEQ ID NO: 2131

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2131

AAGCMATTGT TGTAATTTT GAAAG

25

2) INFORMATION FOR SEQ ID NO: 2132

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2132

TCATATCCAT AGCAATAGTT CTA

23

2) INFORMATION FOR SEQ ID NO: 2133

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: Nucleic acid

1100

- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2133

CCACGCACAW AAACCTTGTTT TAGAAGTAGC AGCWCAGCGT GG

42

2) INFORMATION FOR SEQ ID NO: 2134

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2134

CGCGTGAAGC TTCTGTG

17

2) INFORMATION FOR SEQ ID NO: 2135

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2135

TCTCGCCGTT ATTCAGTTTC

20

2) INFORMATION FOR SEQ ID NO: 2136

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2136

CCAACGCGTC CTCAATCATT TCTAACTTCT ATGGCCGGCG TTGG

44

2) INFORMATION FOR SEQ ID NO: 2137

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 889 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Pseudomonas putida*
- (C) ACCESSION NUMBER: Genome project

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2137

| | | | | | |
|------------|------------|-------------|------------|-------------|-----|
| ATGATCACCG | GTGCTGCCCA | GATGGACGGC | GCGATCCTGG | TTTGCTCGGC | 50 |
| CGCCGATGGT | CCGATGCCrC | AAACCCGTGA | GCACATCCTG | CTGTCCCCTC | 100 |
| AGGTAGGCGT | TCCGTACATC | GTGGTCTTCC | TGAACAAGGC | CGACCTGGTA | 150 |
| GACGACGCTG | AGCTGCTGGA | ACTGGTCGAG | ATGGAAGTTC | GCGACCTGCT | 200 |
| GTCCACCTAC | GACTTCCCAG | GCGACGACAC | TCCGATCATC | ATCGGTTCGG | 250 |
| CTCGTATGGC | CCTGGAAGGC | AAAGACGACA | ACGAAATGGG | CACTACCGCT | 300 |
| GTCAAGAAGC | TGGTAGAAAC | TCTGGATGCC | TACATCCCTG | AGCCAGTTTCG | 350 |
| TGCCATCGAC | CAGCCGTTCC | TGATGCCGAT | CGAAGACGTG | TTCTCGATCT | 400 |
| CGGGTCGTGG | TACCGTTGTT | ACCGGTCGTA | TGAGCGTGG | CATCGTTCGC | 450 |
| GTTTCAGATC | CGCTGGAAAT | CGTTGGTCTG | CGTGACACCA | CCACCACCAC | 500 |
| CTGCACCGGT | GTTGAGATGT | TCCGCAAGCT | GCTGGACGAA | GGTCGTGCTG | 550 |
| GCGAGAACTG | CGGCGTTCTG | CTGCGTGGTA | CCAAGCGTGA | CGACGTTGAG | 600 |
| CGTGGCCAGG | TTCTGGTCAA | GCCAGGTTTCG | GTCAAGCCGC | ACACCAAGTT | 650 |
| CACCGCAGAA | GTCTACGTCC | TGTCGAAGGA | AGAAGGCGGC | CGTCACACTC | 700 |
| CGTTCTTCAA | AGGCTACCGT | CCACAGTTCT | ACTTCCGTAC | CACTGACGTG | 750 |
| ACCGGTAAC | GCGAACTGCC | GGAAGGCGTT | GAAATGGTAA | TGCCAGGTGA | 800 |
| CAACATTGAG | ATGACTGTGA | CCCTGATCAA | GACCATCGCA | ATGGAAGACG | 850 |
| GTCTGCGCTT | CGCTATCCGT | GAAGGCGGTC | GTACCGTCG | | 889 |

2) INFORMATION FOR SEQ ID NO: 2138

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Pseudomonas aeruginosa*
- (B) STRAIN: PAO-1
- (C) ACCESSION NUMBER: Genome project

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2138

| | | | | | |
|-------------|------------|------------|------------|------------|-----|
| AACATGATCA | CCGGTGCTGC | CCAGATGGAC | GGCGCGATCC | TGGTTTGCTC | 50 |
| GGCTGCCGAC | GGCCCCATGC | CGCAGACCCG | CGAGCACATC | CTGCTGTCCC | 100 |
| GCCAGGTAGG | CGTTCCCTAC | ATCGTCGTGT | TCCTGAACAA | GGCCGACATG | 150 |
| GTCGACGACG | CCGAGCTGCT | GGAAGTGGTC | GAGATGGAAG | TTCGCGATCT | 200 |
| GCTGAACACC | TACGACTTCC | CGGGCGACGA | CACTCCGATC | ATCATCGGTT | 250 |
| CCGCGCTGAT | GGCGCTGGAA | GGCAAGGATG | ACAACGGCAT | CGGCGTAAGC | 300 |
| GCCGTGCAGA | AGCTGGTAGA | GACCCTGGAC | TCCTACATTC | CGGAGCCGGT | 350 |
| TCGTGCCATC | GACCAGCCGT | TCCTGATGCC | GATCGAAGAC | GTGTTCTCGA | 400 |
| TCTCCGGCCG | CGGTACCGTG | GTAACCGGTC | GTGTAGAGCG | CGGCATCATC | 450 |
| AAGGTCCAGG | AAGAAGTGGA | AATCGTCGGC | ATCAAGGCGA | CCACCAAGAC | 500 |
| TACCTGCACC | GGCGTTGAAA | TGTTCCGCAA | GCTGCTCGAC | GAAGGTCGTG | 550 |
| CTGGTGAGAA | CGTTGGTATC | CTGCTGCGTG | GCACCAAGCG | TGAAGACGTA | 600 |
| GAGCGTGGCC | AGGTTCTGGC | CAAGCCGGGC | ACCATCAAGC | CGCACACCAA | 650 |
| GTTTCGAGTGC | GAAGTGTAAG | TGCTGTCCAA | GGAAGAAGGT | GGTCGTCACA | 700 |
| CCCCGTTCTT | CAAGGGCTAC | CGTCCGCACT | TCTACTTCCG | TACCACCGAC | 750 |
| GTGACCGGTA | ACTGCGAACT | GCCGGAAGGC | GTAAGAGATG | TAATGCCGGG | 800 |
| CGACAACATC | AAGATGGTTG | TCACCCTGAT | CGCTCCGATC | GCCATGGAAG | 850 |
| ATGGCCTGCG | CTTCGCGATC | CGCGAAGGCG | GCCGTACCGT | TGGCGCCGGC | 900 |

2) INFORMATION FOR SEQ ID NO: 2139

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1398 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Campylobacter jejuni*
- (B) STRAIN: NCTC 11168
- (C) ACCESSION NUMBER: AL139074

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2139

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGCAAGGAT | TTATTTTACA | GGTATTAGGT | CCGGTTGTTG | ATGTAGATTT | 50 |
| TAACGACTAT | TTGCCTCAAA | TTAATGAAGC | AATTGTTGTA | AATTTTGAAA | 100 |
| GCGAAGGAAA | AAAACATAAA | CTTGTTTTAG | AAGTAGCAGC | TCATTTAGGA | 150 |
| GATAATAGAG | TTAGAACTAT | TGCTATGGAT | ATGACAGATG | GTTTGGTAAG | 200 |
| GGGCTTAAAA | GCTGAGGCTT | TAGGTGCTCC | CATTAGTGTT | CCTGTTGGTG | 250 |
| AGAAAGTTTT | AGGAAGAATT | TTCAATGTTA | CTGGAGATTT | GATCGATGAA | 300 |
| GGTGAAGAAA | TTTCTTTTCA | TAAAAAATGG | GCAATTCATA | GAGATCCACC | 350 |
| AGCTTTTGAA | GATCAAAGCA | CAAAAAGTGA | GATTTTTGAA | ACAGGGATTA | 400 |
| AAGTTGTAGA | TTTACTTGCT | CCTTATGCAA | AAGGTGGTAA | AGTAGGTCTT | 450 |
| TTTGGTGGTG | CAGGTGTTGG | TAAAACTGTT | ATTATTATGG | AGCTTATTCA | 500 |
| CAATGTTGCA | TTTAAGCATA | GCGGCTATTC | TGTATTTGCA | GGTGTGGGTG | 550 |
| AGAGAACTCG | TGAAGGAAAT | GACCTTTATA | ATGAAATGAA | AGAAAGTAAT | 600 |
| GTTTTAGACA | AAGTTGCTCT | ATGTTATGGA | CAAATGAATG | AACCACCAGG | 650 |
| GGCAAGAAAT | CGTATTGCTT | TAACAGGTTT | AACAATGGCT | GAGTATTTTA | 700 |
| GAGATGAAAT | GGGTCTTGAT | GTGCTTATGT | TTATTGATAA | TATCTTTAGA | 750 |
| TTTTTACAAT | CAGGTTCTGA | AATGTCAGCA | CTTTTAGGAA | GAATTCCATC | 800 |
| AGCTGTGGGT | TATCAACCAA | CCTTAGCAAG | TGAAATGGGT | AAATTCCAAG | 850 |
| AAAGAATTAC | TTCAACTAAA | AAAGGCTCAA | TCACTTCAGT | TCAAGCTGTT | 900 |
| TACGTTCCAG | CTGATGACTT | AACAGACCCA | GCTCCAGCAA | CTGTTTTTGC | 950 |

| | | | | | |
|------------|-------------|------------|-------------|-------------|------|
| TCACCTAGAT | GCTACAACGG | TTTTAAATAG | AGCTATTGCT | GAAAAGGGTA | 1000 |
| TTTATCCTGC | AGTTGACCCA | CTTGACTCAA | CTTCAAGAAT | GCTTGACCCA | 1050 |
| AATATCATTG | GAGAAGAACA | TTATAAAGTT | GCTCGTGGTG | TTCAATCAGT | 1100 |
| GCTTCAAAAA | TACAAAAGATT | TACAAGATAT | CATCGCCATT | TTAGGTATGG | 1150 |
| ATGAGCTTAG | CGAAGAGGAT | AAACTTGTAG | TTGAAAGAGC | AAGAAAAGATT | 1200 |
| GAAAAATTCT | TATCACAACC | ATTTTTCGTT | GCAGAAGTTT | TCACAGGTAG | 1250 |
| CCCAGGAAAA | TATATAAGCC | TTGAAGATAC | AATAGCAGGA | TTTAAGGGAA | 1300 |
| TTTTAGAAGG | TAAATATGAT | CATTTACCAG | AAAAATGCTTT | CTATATGGTT | 1350 |
| GGAAATATAG | ATGAAGCTAT | TGCAAAAGCG | GATAAATTAA | AAGGTTAA | 1398 |

2) INFORMATION FOR SEQ ID NO: 2140

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2157 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: 670
- (C) ACCESSION NUMBER: X67867

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2140

| | | | | | |
|------------|------------|------------|------------|-------------|------|
| ATGAACAAAC | CAACGATTCT | GCGCCTAATC | AAGTATCTGA | GCATTAGCTT | 50 |
| CTTAAGCTTG | GTTATCGCAG | CCATTGTCTT | AGGCGGAGGA | GTTTTTTTCT | 100 |
| ACTACGTTAG | CAAGGCTCCT | AGCCTATCCG | AGAGTAAACT | AGTTGCAACA | 150 |
| ACTTCTAGTA | AAATCTACGA | CAATAAAAAT | CAACTCATTG | CTGACTTGGG | 200 |
| TTCTGAACGC | CGCGTCAATG | CCCAAGCTAA | TGATATTCCC | ACAGATTTGG | 250 |
| TTAAGGCAAT | CGTTTCTATC | GAAGACCATC | GCTTCTTCGA | CCACAGGGGG | 300 |
| ATTGATTCCA | TCCGTATCCT | GGGAGCTTTC | TTGCGCAATC | TGCAAAGTAA | 350 |
| TTCCCTCCAA | GGTGGATCAA | CTCTCACCCA | ACAGTTGATT | AAGTTGACTT | 400 |
| ACTTTTCAAC | CTCGACTTCC | GACCAGACTA | TTTCTCGTAA | GGCTCAGGAA | 450 |
| GCTTGCTTAG | CGATTCAATT | AGAACAACAA | GCAACCAAAC | AGGAAATCTT | 500 |
| GACCTACTAT | ATAAATAAGG | TCTACATGTC | TAATGGCAAC | TATGGAATGC | 550 |
| AGACAGCAGC | TCAAAACTAC | TATGGTAAAG | ACCTCAATAA | TTTAAGTTTA | 600 |
| CCTCAGTTAG | CCTTGCTGGC | TGGAATGCCT | CAGGCACCAA | ACCAATATGA | 650 |
| CCCCTATTCA | CATCCAGAAG | CAGCCCAAGA | CCGCCGAAAC | TTGGTCTTAT | 700 |
| CTGAAATGAA | AAATCAAGGT | TACATCTCTG | CTGAACAGTA | TGAGAAAGCA | 750 |
| GTCAATACAC | CAATTACTGA | TGGACTACAA | AGTCTCAAAT | CAGCAAGTAA | 800 |
| TTACCCTGCT | TACATGGATA | ATTACCTCAA | GGAGGTCATC | AATCAAGTAG | 850 |
| AACAAGAAAC | TGGCTATAAC | CTTCTAACTA | CTGGGATGGA | TGTTTACACA | 900 |
| AATGTAGACC | AAGAAGCTCA | AAAACATCTG | TGGGATATCT | ACAACCTCCGA | 950 |
| TCAATACGTC | TCTTACCCTG | ACGATGATTT | GCAAGTCGCA | TCTACGGTCG | 1000 |
| TAGATGTTTC | AAATGGTAAA | GTCATCGCCC | AACTTGGAGC | TCGTACCCAA | 1050 |
| GCAAGTAACG | TTTCATTGTT | TACCAACCAA | GCTGTGGAAA | CCAATCGTGA | 1100 |
| CTGGGGTTCT | GCTATGAAAC | CAATCACCGA | TTATGCACCT | GCCATAGAAT | 1150 |
| ACGGTGTTTA | TGATTCCACT | GCAACTATGG | TTAATGATAT | TCCTTATAAC | 1200 |
| TATCCGGGAA | CAAGCACACC | TGTCTACAAC | TGGGATAGAG | CATATTTTCGG | 1250 |
| TAATATTACT | CTGCAATATG | CTCTTCAACA | ATCACGAAAT | GTCACAGCCG | 1300 |
| TTGAGACTTT | GAATAAGGTC | GGTCTAGATA | GAGCTAAAAC | CTTCCTTAAT | 1350 |
| GGTCTTGCTA | TCGACTATCC | AAGCATGCAT | TATGCAAACG | CCATTTC AAG | 1400 |

| | | | | | |
|------------|------------|------------|------------|-------------|------|
| TAATACAACA | GAATCTAATA | AACAATACGG | AGCAAGTAGT | GAAAAAATGG | 1450 |
| CTGCTGCTTA | TGCTGCCTTT | GCCAACGGTG | GAATTTACCA | CAAACCCATG | 1500 |
| TATATCAATA | AGGTCGTCTT | CAGTGACGGT | AGTAAAAAAG | AATTTTCAGA | 1550 |
| TGTAGGTACA | CGAGCTATGA | AAGAAACAAC | TGCTTACATG | ATGACCGAAA | 1600 |
| TGATGAAAAC | TGTCTTGGA | TACGGAAGTG | GTCGTGGAGC | CTATCTCCCA | 1650 |
| TGGTTAGCGC | AAGCTGGTAA | GACAGGTACT | TCTAACTACA | CAGATGATGA | 1700 |
| AATTGAAAAA | CACATCAAGA | AACTGGGCTA | TGTAGCTCCA | GATGAAATGT | 1750 |
| TTGTTGGTTA | TACTCGTAAG | TATTCTATGG | CTGTATGGAC | AGGTTATTTCG | 1800 |
| AATCGTTTAA | CTCCTATCGT | TGGAGATGGT | TTCCTAGTTG | CAGCTAAAGT | 1850 |
| TTATCGCTCA | ATGATAACGT | ATCTATCAGA | AGATACTCAT | CCAGAAGACT | 1900 |
| GGACGATGCC | AGACGGACTT | TTCAGAAACG | GGGAATTTGT | ATTCAAAAAT | 1950 |
| GGAGCTCGTT | CTACGTGGAG | CTCACCTGCT | CCACAACAAC | CCCCATCAAC | 2000 |
| TGAAAGTTCA | AGCTCATCAT | CAGATAGTTC | AACTTCACAG | TCTAGCTCAA | 2050 |
| CCACTCCAAG | CACAAATAAT | AGTACGACTA | CCGATCCTAA | CAATAATACG | 2100 |
| CAACAATCAA | ATACAACCCC | TGATCAACAA | AATCAGAATC | CTCAACCAGC | 2150 |
| ACAACCA | | | | | 2157 |

2) INFORMATION FOR SEQ ID NO: 2141

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2141

CACCAGTCAC AGAAAAGC

18

2) INFORMATION FOR SEQ ID NO: 2142

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2142

GCCTTAATTT CGGATAGTGC

20

2) INFORMATION FOR SEQ ID NO: 2143

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1105

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2143

GAGAAAGAGC CCAATTATCT AATGT

25

2) INFORMATION FOR SEQ ID NO: 2144

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3075 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus aureus*
- (B) STRAIN: J2870
- (C) ACCESSION NUMBER: X75439

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2144

| | | | | | |
|------------|------------|------------|------------|------------|------|
| TTGACAAAGA | AATATTTTAA | CACCCAGAAT | GAAATATCAG | CATTTTGGAA | 50 |
| TACTCAAAAG | ATATTTTAAA | AATCAATTGA | CAATAGAAAA | GGACAGGAAA | 100 |
| GTTTTGTTTT | TTATGACGGC | CCCCCACTG | CAAATGGCCT | TCCTCATGCT | 150 |
| GGCCATGTTT | TTGGAAGAGT | AATCAAGGAT | TTAGTTGCAA | GATTAAAAAC | 200 |
| TATGCAAGGT | TTTTATGTAG | AAAGAAAAGC | AGGATGGGAT | ACCCATGGCT | 250 |
| TACCAGTTGA | ATTAGAGGTT | GAAAAAATAA | TTGGAATTAA | AGGAAAACAA | 300 |
| GACATTGAAA | AGTATGGAAT | AGAAAATTTT | ATAAATGAAT | GTAAAAAAG | 350 |
| TGTATTTAAT | TATGAAAAAG | AATGGCGGGA | TTTTTCTAAA | GATTTAGGAT | 400 |
| ACTGGGTTGA | CATGGACTCC | CCCTATATAA | CTCTTGAGAA | TAATTATATT | 450 |
| GAAAGTGAT | GGAATATATT | ATCTACATTC | CATAAAAAAG | GACTATTATA | 500 |
| TAAGGGACAT | AAGGTGACTC | CTTATTGTAC | ACATGATCAA | ACCGCTTTAA | 550 |
| GTTCTCATGA | AGTAGCGCAA | GGCTATAAAA | ACGTTAAAGA | TTTATCAGCT | 600 |
| GTTGTTAAT | TTCAACTTAC | AAATAGTAAA | GATACTTATT | TCTTAAGTTG | 650 |
| GACTACCACT | CCCTGGACTT | TGCCTGCAAA | TGTAGCATT | GCTATAAATA | 700 |
| AAGATCTTAA | TTATTCAAAA | ATTCGGGTAG | AAAATGAGTA | TTATATCTTA | 750 |
| GCTACAGATC | TAATTAATTC | TATAATAACT | GAAAAATACG | AAATTATTGA | 800 |
| TACCTTTTCA | GGAAGTAATT | TAATTAATTT | AAAATACATT | CCTCCTTTTG | 850 |
| AAAGCGACGG | TTTAGTTAAT | GCATATTACG | TTGTTGATGG | AGAATTTGTT | 900 |
| ACTAACTCAG | AAGGAACTGG | TATTGTTTAT | ATAGCACCAG | CTCATGGGGA | 950 |
| AGATGACTAC | CAATTGGTTT | TAGAGCGTGA | TTTGGAATTC | TTAAATGTTA | 1000 |
| TAACAAGAGA | AGGAGTATAT | AATGATAGGT | TCCCTGAATT | AGTTGGTAAT | 1050 |
| AAAGCTAAAA | ATAGTGATAT | AGAAATCATA | AAATTATTAT | CCAAAAACA | 1100 |
| ACTTTTATAT | AAAAAACAAA | AATATGAGCA | TAATTATCCT | CATTGTTGGA | 1150 |
| GATGTGGTAA | TCCTTTGATA | TATTATGCGA | TGGAAGGTTG | GTTTATTAAA | 1200 |
| ACAATAATT | TTAAGAATGA | AATTATTAAC | AATAATAATA | ATATAGAGTG | 1250 |
| GTTTCCTTCT | CATATTAAGG | AAGGGAGAAT | GGGAAATTTT | TTAGAAAATA | 1300 |
| TGGTTGATTG | GAACATTGGT | AGAAATAGAT | ATTGGGGAAC | ACCATTAAAT | 1350 |
| GTATGGATTT | GCAATGATTG | TAATCACGAA | TACGCACCAA | GTAGTATTAA | 1400 |
| GGATTTACAA | AATAATTCCA | TCAATAAAAT | TGATGAAGAT | ATTGAGTTGC | 1450 |
| ATAGACCTTA | TGTTGATAAT | ATCACTCTTA | GTTGCCCTAA | GTGTAATGGG | 1500 |

1106

| | | | | | |
|-------------|-------------|-------------|------------|------------|------|
| AAAATGTCTC | GAGTAGAAGA | AGTAATCGAT | GTTTGGTTTG | ATAGCGGCTC | 1550 |
| TATGCCGTTT | GCTCAGCATC | ATTATCCTTT | TGATAACCAG | AAAATTTTAA | 1600 |
| ATCAACACTT | TCCAGCTGAT | TTTATTGCAG | AAGGAGTTGA | TCAAACGAGA | 1650 |
| GGCTGGTTTT | ACAGTTTACT | AGTAATTTCT | ACTATTCTAA | AAGGAAAATC | 1700 |
| TTCTTATAAA | CGTGCTTTAT | CTTTAGGACA | TATTCTAGAC | AGTAATGGTA | 1750 |
| AAAAAATGTC | TAAAAGTAAA | GGAAACGTTA | TTAATCCAAC | TGAATTAATT | 1800 |
| AATAAGTACG | GAGCCGATTTC | TTTAAGATGG | GCCTTAATTT | CGGATAGTGC | 1850 |
| TCCATGGAAT | AACAAAAGAT | TCTCAGAAAA | TATAGTAGCT | CAGACCAAAT | 1900 |
| CGAAATTTAT | AGATACGCTT | GATAATATTT | ATAAATTTTA | TAATATGTAT | 1950 |
| AATAAAATAG | ATCACTATAA | TCCTAATAAT | GAAATTACAA | AAAGTAGAAA | 2000 |
| TACATTAGAT | AATTGGGCTC | TTTCTCGCTT | AAACACCTTA | ATAAAAGAAA | 2050 |
| GTAATATTTA | TGTAAATAAT | TACGATTTCA | CTTCCGCAGC | CAGATTAATT | 2100 |
| AACGAATATA | CCAATACAAT | AAGTAATTGG | TATATCGGAG | ATTCGAGAGG | 2150 |
| ACGATTTTGG | GAACAAGGAA | TTTCTAACGA | TAAAAAAGAT | GCGTACAATA | 2200 |
| CGCTTTATGA | AATTTTAACA | ACTTTATCAA | GACTAGTGGC | TCCATTTGTT | 2250 |
| CCATTTATAT | CTGAAAAAAT | CCATTATAAT | TTGACTGGAA | AAAGTGTGCA | 2300 |
| TTTACAAGAT | TATCCACAAT | ATAAAGAAAG | TTTTATTAAT | CAAGCATTGG | 2350 |
| AAGATGAAAT | GCATACCGTT | ATAAAAATTG | TAGAATTATC | TAGACAGGCT | 2400 |
| CGCAAAAATG | CAGATTTAAA | AATTAAGCAA | CCTTTATCGA | AAATGGTGAT | 2450 |
| TAAACCTAAT | AGTCAATTAA | ACTTAAGTTT | TTTACCTAAT | TACTATTCAA | 2500 |
| TAATAAAAGA | CGAATTAAAT | ATAAAAAACA | TTGAATTAAC | TGATAATATT | 2550 |
| AATGACTATA | TTACCTATGA | GCTTAAATTG | AATTTTTCTT | CTGTGGGACC | 2600 |
| AAAAC TAGGG | AACAAAACGA | AAAATATTCA | AACATTGATA | GACTCCCTAT | 2650 |
| CAGAGTATGA | TAAAAAAAGT | TTAATTGAGT | CTAATAACTT | CAAAAGTTTA | 2700 |
| TCTTCTGATG | CTGAGTTAAC | TAAGGATGAT | TTTATAATTA | AAACCTTACC | 2750 |
| TAAGGATAGT | TATCAACTCA | GTGAAGATAA | TGACTGCGTT | ATATTATTAG | 2800 |
| ATAAAAATTT | ATCTCCTGAA | TTAATTCGCG | AAGGACATGC | TAGAGAGCTC | 2850 |
| ATTAGATTAA | TTCAACAATT | AAGAAAAAAG | AAAAATTTAC | CAATAAATCA | 2900 |
| ACGTATTGAT | ATTTATATCG | GTGTAAC TGG | GGAATTATTA | GAATCAATAA | 2950 |
| AAACCAATAA | AAATATGTTT | AAAGAAAATT | TCGTGATTAA | AAATATACAC | 3000 |
| TTAAATGTTA | TAGATGAATA | TGAAAATACT | ATTCATTTTA | ATAATAAAGA | 3050 |
| AATAAAAATT | TCCTTATTAT | ATTAA | | | 3075 |

2) INFORMATION FOR SEQ ID NO: 2145

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2145

GCAAGATGTG GCGTGTTACG GT

22

2) INFORMATION FOR SEQ ID NO: 2146

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1107

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2146

GGGGCGAAGA AGTTGTCCAT ATT

23

2) INFORMATION FOR SEQ ID NO: 2147

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 660 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (C) ACCESSION NUMBER: V00622

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2147

| | | | | | |
|-------------|------------|-------------|------------|-------------|-----|
| ATGGAGAAAA | AAATCACTGG | ATATACCACC | GTTGATATAT | CCCAATGGCA | 50 |
| TCGTAAAGAA | CATTTTGAGG | CATTTCAGTC | AGTTGCTCAA | TGTACCTATA | 100 |
| ACCAGACCGT | TCAGCTGGAT | ATTACGGCCT | TTTTAAAGAC | CGTAAAGAAA | 150 |
| AATAAGCACA | AGTTTATCC | GGCCTTTATT | CACATTCTTG | CCCGCCTGAT | 200 |
| GAATGCTCAT | CCGGAATTCC | GTATGGCAAT | GAAAGACGGT | GAGCTGGTGA | 250 |
| TATGGGATAG | TGTTACCCCT | TGTTACACCG | TTTTCCATGA | GCAAACCTGAA | 300 |
| ACGTTTTTCAT | CGCTCTGGAG | TGAATACCAC | GACGATTTCC | GGCAGTTTCT | 350 |
| ACACATATAT | TCGCAAGATG | TGGCGTGTTA | CGGTGAAAAC | CTGGCCTATT | 400 |
| TCCCTAAAGG | GTTTATTGAG | AATATGTTTT | TCGTCTCAGC | CAATCCCTGG | 450 |
| GTGAGTTTCA | CCAGTTTGA | TTTAAACGTG | GCCAATATGG | ACAACTTCTT | 500 |
| CGCCCCCGTT | TTCACCATGG | GCAAATATTA | TACGCAAGGC | GACAAGGTGC | 550 |
| TGATGCCGCT | GGCGATTCAG | GTTTCATCATG | CCGTCTGTGA | TGGCTTCCAT | 600 |
| GTCGGCAGAA | TGCTTAATGA | ATTACAACAG | TACTGCGATG | AGTGGCAGGG | 650 |
| CGGGGCGTAA | | | | | 660 |

2) INFORMATION FOR SEQ ID NO: 2148

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2148

CAGATTAAAT GCGGATTCAG CC

22

2) INFORMATION FOR SEQ ID NO: 2149

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2149

ATCAGGTAAA TCATCAGCGG ATA

23

2) INFORMATION FOR SEQ ID NO: 2150

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 642 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: K12
- (C) ACCESSION NUMBER: X53796

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2150

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGAATTTTA | CCCGGATTGA | CCTGAATACC | TGGAATCGCA | GGGAACACTT | 50 |
| TGCCCTTTAT | CGTCAGCAGA | TTAAATGCGG | ATTGAGCCTG | ACCACCAAAC | 100 |
| TCGATATTAC | CGCTTTGCGT | ACCGCACTGG | CGGAGACAGG | TTATAAGTTT | 150 |
| TATCCGCTGA | TGATTTACCT | GATCTCCCGG | GCTGTTAATC | AGTTTCCGGA | 200 |
| GTTCCGGATG | GCACTGAAAG | ACAATGAACT | TATTTACTGG | GACCAGTCAG | 250 |
| ACCCGGTCTT | TACTGTCTTT | CATAAAGAAA | CCGAAACATT | CTCTGCACTG | 300 |
| TCCTGCCGTT | ATTTTCCGGA | TCTCAGTGAG | TTTATGGCAG | GTTATAATGC | 350 |
| GGTAACGGCA | GAATATCAGC | ATGATACCAG | ATTGTTTCCG | CAGGGAAATT | 400 |
| TACCGGAGAA | TCACCTGAAT | ATATCATCAT | TACCGTGGGT | GAGTTTTGAC | 450 |
| GGATTTAACC | TGAACATCAC | CGGAAATGAT | GATTATTTTG | CCCCGGTTTT | 500 |
| TACGATGGCA | AAGTTTCAGC | AGGAAGGTGA | CCGCGTATTA | TTACCTGTTT | 550 |
| CTGTACAGGT | TCATCATGCA | GTATGTGATG | GCTTTCATGC | AGCACGGTTT | 600 |
| ATTAATACAC | TTCAGCTGAT | GTGTGATAAC | ATACTGAAAT | AA | 642 |

2) INFORMATION FOR SEQ ID NO: 2151

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid

1109

- (C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2151

ATATTTCAGC ATTACCTTGG GTT

23

2) INFORMATION FOR SEQ ID NO: 2152

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2152

TACACAAC TC TTGTAGCCGA TTA

23

2) INFORMATION FOR SEQ ID NO: 2153

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 642 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Shigella flexineri*
(C) ACCESSION NUMBER: X07848

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2153

| | | | | | |
|------------|------------|-------------|-------------|------------|-----|
| ATGAACTATA | CAAAATTTGA | TGTAAAAAAT | TGGGTTTCGCC | GTGAGCATTT | 50 |
| TGAGTTTTAT | CGGCATCGTT | TACCATGTGG | TTTTAGCTTA | ACAAGCAAAA | 100 |
| TTGATATCAC | GACGTTAAAA | AAGTCATTGG | ATGATTCAGC | GTATAAGTTT | 150 |
| TATCCGGTAA | TGATCTATCT | GATTGCTCAG | GCCGTGAATC | AATTTGATGA | 200 |
| GTTGAGAATG | GCGATAAAAG | ATGATGAATT | GATCGTATGG | GATTCAGTCG | 250 |
| ACCCACAATT | CACCGTATTC | CATCAAGAAA | CAGAGACATT | TTCAGCACTG | 300 |
| AGTTGCCCAT | ACTCATCCGA | TATTGATCAA | TTTATGGTGA | ATTATTTATC | 350 |
| GGTAATGGAA | CGTTATAAAA | GTGATACCAA | GTTATTTTCCT | CAAGGGGTAA | 400 |
| CACCAGAAAA | TCATTTAAAT | ATTTTCAGCAT | TACCTTGGGT | TAATTTTGAT | 450 |
| AGCTTTAATT | TAAATGTTGC | TAATTTTACC | GATTATTTTG | CACCCATTAT | 500 |
| AACAATGGCA | AAATATCAGC | AAGAAGGGGA | TAGACTGTTA | TTGCCGCTCT | 550 |
| CAGTACAGGT | TCATCATGCA | GTTTGTGATG | GCTTCCATGT | TGCACGCTTT | 600 |
| ATTAATCGGC | TACAAGAGTT | GTGTAACAGT | AAATTAAAAAT | AA | 642 |

2) INFORMATION FOR SEQ ID NO: 2154

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2154

CGCCATTCAG AGTTTAGGAC

20

2) INFORMATION FOR SEQ ID NO: 2155

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2155

TTCCATACCG TTGCGTATCA CTT

23

2) INFORMATION FOR SEQ ID NO: 2156

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 624 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium perfringens*
- (B) STRAIN: CP590
- (C) ACCESSION NUMBER: M74769

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2156

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGGTATTTG | AAAAAATTGA | TAAAAATAGT | TGGAACAGAA | AAGAGTATTT | 50 |
| TGACCACTAC | TTTGCAAGTG | TACCTTGATC | ATACAGCATG | ACCGTTAAAG | 100 |
| TGGATATCAC | ACAAATAAAG | GAAAAGGGAA | TGAAACTATA | TCCTGCAATG | 150 |
| CTTTATTATA | TTGCAATGAT | TGTAAACCGC | CATTGAGAGT | TTAGGACGGC | 200 |
| AATCAATCAA | GATGGTGAAT | TGGGGATATA | TGATGAGATG | ATACCAAGCT | 250 |

1111

| | | | | | |
|-------------|------------|-------------|------------|------------|-----|
| ATACAATATT | TCACAATGAT | ACTGAAACAT | TTTCCAGCCT | TTGGACTGAG | 300 |
| TGTAAGTCTG | ACTTTAAATC | ATTTT TAGCA | GATTATGAAA | GTGATACGCA | 350 |
| ACGGTATGGA | AACAATCATA | GAATGGAAGG | AAAGCCAAAT | GCTCCGGAAA | 400 |
| ACATTTT TAA | TGTATCTATG | ATACCGTGGT | CAACCTTCGA | TGGCTTTAAT | 450 |
| CTGAATTTGC | AGAAAGGATA | TGATTATTTG | ATTCCCTATT | TTACTATGGG | 500 |
| GAAATATTAT | AAAGAAGATA | ACAAAATTAT | ACTTCCTTTG | GCAATTCAAG | 550 |
| TTCATCACGC | AGTATGTGAC | GGATTT CACA | TTTGCCGTTT | TGTAAACGAA | 600 |
| TTGCAGGAAT | TGATAAATAG | TTAA | | | 624 |

2) INFORMATION FOR SEQ ID NO: 2157

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2157

CCACAGAAAT TGATATTAGT GTTTTAT

27

2) INFORMATION FOR SEQ ID NO: 2158

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2158

TCGCTATTGT AACCAGTTCT A

21

2) INFORMATION FOR SEQ ID NO: 2159

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 651 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus aureus*
- (C) ACCESSION NUMBER: V01277

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2159

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGAACTTTA | ATAAAATTGA | TTTAGACAAT | TGGAAGAGAA | AAGAGATATT | 50 |
| TAATCATTAT | TTGAACCAAC | AAACGACTTT | TAGTATAACC | ACAGAAATTG | 100 |
| ATATTAGTGT | TTTATACCGA | AACATAAAAC | AAGAAGGATA | TAAATTTTAC | 150 |
| CCTGCATTTA | TTTTCTTAGT | GACAAGGGTG | ATAAACTCAA | ATACAGCTTT | 200 |
| TAGAACTGGT | TACAATAGCG | ACGGAGAGTT | AGGTTATTGG | GATAAGTTAG | 250 |
| AGCCACTTTA | TACAATTTTT | GATGGTGTAT | CTAAAACATT | CTCTGGTATT | 300 |
| TGGACTCCTG | TAAAGAATGA | CTTCAAAGAG | TTTTATGATT | TATACCTTTC | 350 |
| TGATGTAGAG | AAATATAATG | GTTTCGGGGA | ATTGTTTCCC | AAAACACCTA | 400 |
| TACCTGAAAA | TGCTTTTTCT | CTTTCTATTA | TTCCATGGAC | TTCATTTACT | 450 |
| GGGTTTAACT | TAAATATCAA | TAATAATAGT | AATTACCTTC | TACCCATTAT | 500 |
| TACAGCAGGA | AAATTCATTA | ATAAAGGTAA | TTCAATATAT | TTACCGCTAT | 550 |
| CTTTACAGGT | ACATCATTCT | GTTTGTGATG | GTTATCATGC | AGGATTGTTT | 600 |
| ATGAACTCTA | TTCAGGAATT | GTCAGATAGG | CCTAATGACT | GGCTTTTATA | 650 |
| A | | | | | 651 |

2) INFORMATION FOR SEQ ID NO: 2160

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2160

TTTTGAACAC TATTTTAACC AGC

23

2) INFORMATION FOR SEQ ID NO: 2161

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2161

GATTTAACTT ATCCCAATAA CCT

23

2) INFORMATION FOR SEQ ID NO: 2162

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 648 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double

1113

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Staphylococcus aureus*

(C) ACCESSION NUMBER: X02166

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2162

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGACTTTTA | ATATTATCAA | ATTAGAAAAT | TGGGATAGAA | AAGAATATTT | 50 |
| TGAACACTAT | TTTAACCAGC | AAACTACGTA | TAGCATTACT | AAAGAAATTG | 100 |
| ATATTACTTT | GTTTAAAGAT | ATGATAAAAA | AGAAAGGATA | TGAAATTTAT | 150 |
| CCTTCTTTGA | TTTATGCAAT | TATGGAAGTT | GTAAATAAAA | ATAAAGTGTT | 200 |
| TAGAACAGGA | ATTAATAGTG | AGAATAAATT | AGGTTATTGG | GATAAGTTAA | 250 |
| ATCCTTTGTA | TACAGTTTTT | AATAAGCAAA | CTGAAAAAAT | TACTAACATT | 300 |
| TGGACTGAAT | CTGATAACAA | CTTCACTTCT | TTTTATAATA | ATTATAAAAA | 350 |
| TGACTTGCTT | GAATATAAAG | ATAAAGAAGA | AATGTTTCCT | AAAAAACCGA | 400 |
| TACCTGAAAA | CACCATACCG | ATTTCAATGA | TTCCTTGGAT | TGATTTTAGT | 450 |
| TCATTTAATT | TAAACATTGG | TAACAATAGC | AACTTTTAT | TGCCTATTAT | 500 |
| TACGATAGGT | AAATTTTATA | GTGAGAATAA | TAAAATTTAT | ATACCAGTTG | 550 |
| CTTTGCAGCT | TCATCATGCT | GTATGTGATG | GTTACCATGC | TTCATTATTT | 600 |
| ATGAATGAAT | TTCAAGATAT | AATTCATAAG | GTAGATGATT | GGATTTAG | 648 |

2) INFORMATION FOR SEQ ID NO: 2163

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2163

ACCTTCATCC TACCGATGTG GGTT

24

2) INFORMATION FOR SEQ ID NO: 2164

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2164

CAACGACACC AGCACTGCCA TTG

23

2) INFORMATION FOR SEQ ID NO: 2165

1114

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1215 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella typhimurium*
- (B) STRAIN: H3380
- (C) ACCESSION NUMBER: AF071555

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2165

| | | | | | |
|------------|------------|------------|------------|------------|------|
| ATGACCACCA | CACGCCCCGC | GTGGGCCTAT | ACGCTGCCGG | CAGCACTGCT | 50 |
| GCTGATGGCT | CCTTTCGACA | TCCTCGCTTC | ACTGGCGATG | GATATTTATC | 100 |
| TCCCTGTCGT | TCCAGCGATG | CCCGGCATCC | TGAACACGAC | GCCCGCTATG | 150 |
| ATCCAACTCA | CGTTGAGCCT | CTATATGGTG | ATGCTCGGCG | TGGGCCAAGT | 200 |
| GATTTTTGGT | CCGCTCTCAG | ACAGAATCGG | GCGACGGCCA | ATTCTACTTG | 250 |
| CGGGCGCAAC | GGCTTTCGTC | ATTGCGTCTC | TGGGAGCAGC | TTGGTCTTCA | 300 |
| ACTGCACCGG | CCTTTGTGCG | TTTCCGTCTA | CTTCAAGCAG | TGGGCGCGTC | 350 |
| GGCCATGCTG | GTGGCGACGT | TCGCGACGGT | TCGCGACGTT | TATGCCAACC | 400 |
| GTCCTGAGGG | TGTCGTCATC | TACGGCCTTT | TCAGTTCGAT | GCTGGCGTTC | 450 |
| GTGCCTGCGC | TCGGCCCTAT | CGCCGGAGCA | TTGATCGGCG | AGTTCTTGGG | 500 |
| ATGGCAGGCG | ATATTCATTA | CTTTGGCTAT | ACTGGCGATG | CTCGCACTCC | 550 |
| TAAATGCGGG | TTTCAGGTGG | CACGAAACCC | GCCCTCTGGA | TCAAGTCAAG | 600 |
| ACGCGCCGAT | CTGTCTTGCC | GATCTTCGCG | AGTCCGGCTT | TTTGGGTTTA | 650 |
| CACTGTCGGC | TTTAGCGCCG | GTATGGGCAC | CTTCTTCGTC | TTCTTCTCGA | 700 |
| CGGCTCCCCG | TGTGCTCATA | GGCCAAGCGG | AATATTCCGA | GATCGGATTC | 750 |
| AGCTTTGCCT | TCGCCACTGT | CGCGCTTGTA | ATGATCGTGA | CAACCCGTTT | 800 |
| CGCGAAGTCC | TTTGTCTGTC | GATGGGGCAT | CGCAGGATGC | GTGGCGCGTG | 850 |
| GGATGGCGTT | GCTTGTTTGC | GGAGCGGTCC | TGTTGGGGAT | CGGCGAACTT | 900 |
| TACGGCTCGC | CGTCATTCC | CACCTTCATC | CTACCGATGT | GGGTGTGTCG | 950 |
| GGTCGGTATT | GTCTTCACGG | TGTCCGTTAC | CGCGAACGGC | GCTTTGGCAG | 1000 |
| AGTTCGACGA | CATCGCGGGA | TCAGCGGTGC | CGTTCTACTT | CTGCATCCAA | 1050 |
| AGCCTGATAG | TCAGTATCGT | CGGGACATTG | GCGGTGACGC | TGTTAAACGG | 1100 |
| CGATACAGCG | TGGCCCGTGA | TTTGTTACGC | CACGGCAATG | GCAGTGCTGG | 1150 |
| TGTCGTTGGG | GCTGGCGCTC | CTTCGATCCC | GTGATGCTGC | CACCGAGAAG | 1200 |
| TCGCCAGTCG | TCTAG | | | | 1215 |

2) INFORMATION FOR SEQ ID NO: 2166

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2166

GACAAACCAT TCCTGCTG

2) INFORMATION FOR SEQ ID NO: 2167

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2167

CAGCAGCTGG GCGGCGGT

18

2) INFORMATION FOR SEQ ID NO: 2168

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2168

CATCAAAGTT GGTGAAGAAG TTG

23

2) INFORMATION FOR SEQ ID NO: 2169

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2169

CCCGTTTGCG AAAGGTGG

18

2) INFORMATION FOR SEQ ID NO: 2170

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

1116

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2170

ACGTGACGTT GACAAACCA

19

2) INFORMATION FOR SEQ ID NO: 2171

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2171

TCGTTGGATT AACTGAAGAA

20

2) INFORMATION FOR SEQ ID NO: 2172

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2172

GTGTTGAAAT GTTCCGTAAA

20

2) INFORMATION FOR SEQ ID NO: 2173

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2173

AAGAAAAAAT CTTCGAACTG GCTA

24

2) INFORMATION FOR SEQ ID NO: 2174

1117

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2174

TCTACACGGC CGGTG

15

2) INFORMATION FOR SEQ ID NO: 2175

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2175

CCGCCATACC CCGTTT

16

2) INFORMATION FOR SEQ ID NO: 2176

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2176

CGGCATTACC ATTTCCACAC CTTT

24

2) INFORMATION FOR SEQ ID NO: 2177

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2177

GGCACGGACA AACCATTCTT GCTGCCTATC GAAGACGTGT TCCCGTGCC 49

2) INFORMATION FOR SEQ ID NO: 2178

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2178

GGCACGACAA ACCATTCTTG CTGCCTATCG AACGTGCC 38

2) INFORMATION FOR SEQ ID NO: 2179

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2179

GGCAGCTCTA CTTCCGTACC ACTGACGTAA CCGGCTGCC 39

2) INFORMATION FOR SEQ ID NO: 2180

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2180

TTCGCCGGCG TGGGC 15

2) INFORMATION FOR SEQ ID NO: 2181

(i) SEQUENCE CHARACTERISTICS:

1119

- (A) LENGTH: 15 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2181

AGCGCCACGC GCAGG

15

2) INFORMATION FOR SEQ ID NO: 2182

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2182

GCGGCGCAAC GACTTCTACC ACGAAATGGA AGAGTCGCGC GC

42

2) INFORMATION FOR SEQ ID NO: 2183

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 817 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Alcaligenes faecalis* subsp. *faecalis*
- (B) STRAIN: ATCC 8750

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2183

| | | | | | |
|------------|------------|------------|-------------|-------------|-----|
| TATCTTGGTT | WGCTCGGCCG | CTGACGGCCC | AATGCCTCAG | ACTCGCGAGC | 50 |
| ACATCCTGCT | GAGCCGTCAG | GTTGGCGTTC | CTTACATCAT | CGTGTTTCCTG | 100 |
| AACAAGGCCG | ACATGGTTGA | TGACGAAGAG | CTGATCGAAC | TGGTTGAAAT | 150 |
| GGAAGTTCGC | GAGCTGTTGT | CCAAGTACGA | CTTCCCTGGC | GACGACACCC | 200 |
| CGATCATCAA | GGGTTGCGCC | AAACTGGCTC | TGGAAGGCCA | CGAAGGCCCA | 250 |
| CTGGGCAGCC | AAGCCGTTCT | GGCTCTGGCC | GAAGCGCTGG | ACAACTACAT | 300 |
| TCCTACGCCT | GAGCGTGCCG | TTGACGGTAC | GTTCCCTGATG | CCTGTTGAAG | 350 |
| ACGTGTTCTC | GATCTCCGGC | CGTGGTACGG | TTGTGACCGG | TCGTATTGAG | 400 |
| CGCGGCATCA | TCAAGGTCGG | CGAAGAAATC | GAAATCGTGG | GTATCAAAGA | 450 |
| CACGGTCAAG | ACCATTGTA | CCGGCGTTGA | AATGTTCCGC | AAACTGCTGG | 500 |
| ACCAGGGCGA | AGCTGGCGAT | AACGTCGGTC | TGCTGCTGCG | TGGTACCAAG | 550 |
| CGTGAAGACG | TGGAACGTGG | TCAAGTTCTG | GCCAAGCCAG | GCTCGATCAA | 600 |
| GCCACACACT | GACTTCGACG | CCGAGGTGTA | CATTCTGTCC | AAAGAAGAAG | 650 |
| GTGGTCGTCA | CACTCCTTTC | TTCAAGGGCT | ACCGTCCTCA | GTTCTACTTC | 700 |

| | | | | | | | |
|------------|------------|------------|------------|------------|----------|------------|-----|
| CGTACAAC | CTG | CGG | CACCAT | CGAG | CTGCCAGA | ACAAGGAAAT | 750 |
| GGTTCTGCCA | GGCGACAACA | TTTCGATGAA | AGTGTCCCTG | ATCGCTCCTA | | | 800 |
| TCGCCATGGA | AGAAGGT | | | | | | 817 |

2) INFORMATION FOR SEQ ID NO: 2184

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1652 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Campylobacter coli*
- (B) STRAIN: ATCC 43479

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2184

| | | | | | |
|------------|------------|------------|-------------|------------|------|
| TACAATTGAA | GTTGAAAGAT | CTATGCGTGT | TCTTGATGGT | GCTGTTGCGG | 50 |
| TATTTTGTTC | AGTGGGTGGG | GTGCAGCCTC | AAAGTGAAAC | AGTTTGGAGA | 100 |
| CAAGCAAATA | AATATGGTGT | TCCAAGAATA | GTATTTGTAA | ATAAAATGGA | 150 |
| TAGATCGGT | GCAAATTTCT | ACAATGTAGA | AGATCAAATT | CGCAACCGTT | 200 |
| TAAAAGCTAA | TCCAGTTCCA | CTTCAAATTC | CAATCGGTGC | TGAGGATAAT | 250 |
| TTTAAAGGCG | TAATCGATCT | TGTAACATG | AAAGCTTTAG | TTTGGGAAGA | 300 |
| TGATACTAAG | CCAACGGATT | ATGTAGAAAA | AGAAATTTCCA | GCTGAACTTA | 350 |
| AAGAAAAGGC | AGAAGAATAT | CGCACAAAAA | TGATAGAAGC | AGTTTCTGAA | 400 |
| ACTTCAGATG | AGTTGATGGA | AAAATATTTA | GGTGGAGAAG | AATTAAGCCT | 450 |
| TGAAGAGATT | AAAACAGGGA | TTAAAGCAGG | ATGTTTAAGT | CTTCTATCG | 500 |
| TTCCTATGCT | TTGCGGTACA | GCGTTTAAAA | ATAAAGGGGT | TCAACCTTTG | 550 |
| CTTGATGCTG | TTGTGGCTTA | TTTACCAGCT | CCTGATGAAG | TTGCTAATAT | 600 |
| CAAAGGGGAA | TATGAAGACG | GCACAGAAGT | TTCTGTAAAA | TCAACTGATG | 650 |
| ATGGCGAATT | TGCAGGACTT | GCATTTAAAA | TTATGACAGA | TCCATTTGTA | 700 |
| GGACAACTTA | CTTTCGTGCG | TGTTTATCGT | GGATGTTTAG | AAAGCGGTTT | 750 |
| TTATGCTTAT | AACTCAACCA | AAGATAAAAA | AGAAAGAATT | GGTCGTTTGT | 800 |
| TAAAAATGCA | CTCTAATAAA | AGAGAAGAAA | TTAAAGTTCT | TTACGCAGGA | 850 |
| GAAATCGGTG | CAGTTGTAGG | ACTTAAAGAT | ACTTTAACAG | GGGATACTCT | 900 |
| TGCAAGTGAA | AAAGATAAAG | TAATTCTTGA | AAGAATGGAT | TTCCCAGACC | 950 |
| CAGTTATTTT | TGTTGCAGTA | GAMCCAAAAA | CTAAAGCAGA | TCAAGAAAAA | 1000 |
| ATGTCTATTG | CACTAAATAA | ATTAGCTCAA | GAAGATCCAA | GCTTTAGAGT | 1050 |
| TTCTACGGAT | GAAGAAAGTG | GTCAAACAT | CATTTACAGG | ATGGGTGAAT | 1100 |
| TACACCTTGA | AATTATTGTT | GATCGTATGC | TTCTGTGAAT | TAAAGTGGAA | 1150 |
| GCTGAAGTGG | GACAACCTCA | AGTTGCTTAT | CGTGAAACTA | TCAGAAAAAC | 1200 |
| TGTTGAGCAA | GAATACAAAT | ACGCTAAGCA | ATCAGGCGGT | CGTGGTCAGT | 1250 |
| ATGGACATGT | ATTCTTACGC | CTTGAGCCAC | TTGAGCCAGG | AAGTGGATAC | 1300 |
| GAGTTTGTTA | ACGACATCAA | AGGTGGGGTA | ATTCCAAAAG | AATATATTCC | 1350 |
| TGCAGTAGAT | AAGGGTGTTT | AAGAAGCATT | GCAAAATGGT | GTTTTAGCAG | 1400 |
| GTTATCCAGT | AGAAGATGTT | AAAGTAACTG | TTTATGATGG | AAGTTATCAC | 1450 |
| GAGGTGGATT | CATCTGAGAT | GGCATTTAAA | CTTGCTGCTT | CTATGGGATT | 1500 |
| TAAAGAGGGT | GCTAGAAAAG | CAGGTGCTGT | GATCTTAGAG | CCTATGATGA | 1550 |
| AAGTTGAAGT | AGAACTCCT | GAAGATTACA | TGGGCGATGT | TATTGGCGAT | 1600 |
| CTTAATAAGC | GTCGTGGTCA | AGTAAATAGC | ATGGATGAAA | GAGGTGGTAA | 1650 |
| TA | | | | | 1652 |

2) INFORMATION FOR SEQ ID NO: 2185

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 820 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Succinivibrio dextrinosolvens*
- (B) STRAIN: ATCC 19716

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2185

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GCTATTCTAG | TAGTAGCAGC | AACTGATGGT | CCTATGCCAC | AGACCCGTGA | 50 |
| GCACATCCTA | TTAGCACGTC | AGGTAGGCGT | ACCATACATC | ATCGTATTCC | 100 |
| TAAACAAGTG | CGATATGGTT | GACGACGAGG | AATTATTAGA | GTTAGTTGAG | 150 |
| ATGGACGTAC | GTGATCTATT | AAATCAGTAC | CAGTTCCCAG | GCGACGACAC | 200 |
| TCCAATCATC | CGTGGTTCAG | CACTAGGTGC | ATTAAACGGC | GAAGAGAAGT | 250 |
| GGAAAGAGGC | AATCTATCAG | TTAGCAGACA | CTCTAGATTG | ATACATTCCA | 300 |
| GAGCCAAAGC | GTGATATCGA | TGATCCATTG | CTATTACCAA | TCGAAGATAT | 350 |
| CTTCTCAATC | TCAGGTCGTG | GTACTGTAGT | AACCGGCCGT | GTAGAGCGTG | 400 |
| GTATTGTACA | CGTAGGTGAC | GAAGTTGAAA | TCGTTGGTAT | TCGTCCAACC | 450 |
| ACCAAGACCA | CTGTAAGTGG | CGTTGAAATG | TTCCGTAAGT | TACTAGACGA | 500 |
| AGGTCGTGCA | GGTGATAACG | TTGGTGTTCT | ACTACGTGGT | ACCAAGCGTG | 550 |
| ATGAGGTTGA | GCGTGGTCAG | GTTCTAGCTG | CTCCAGGCAC | AATCACTCCA | 600 |
| CACACCAAGT | TCACTGGTCA | GGTTTACGTA | CTAAGCAAGG | ATGAAGGTGG | 650 |
| TCGTCACACT | CCATTCTTCA | AGGGCTACCG | TCCACAGTTC | TTCTTCCGTA | 700 |
| CAACCGATAT | TACCGGTTCT | ATCGATCTGA | AAGAGGGCGT | AGAGATGGTA | 750 |
| ATGCCAGGTG | ATAACACCGA | CATGACCGTA | ACCCTAATCC | ACCCAGTAGC | 800 |
| TATGGCTGAA | GGCGAGAGAT | | | | 820 |

2) INFORMATION FOR SEQ ID NO: 2186

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2186

ACGCGCTCAA AGCAGAAGTA TACGTATTAT CAAAAGACGC GCGT

44

2) INFORMATION FOR SEQ ID NO: 2187

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1612 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Campylobacter jejuni* subsp. *jejuni*
- (B) STRAIN: ATCC 33292

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2187

```

CAGTGGGTGG GGTGCAGCCT CAAAGTGAAA CAGTTTGGAG ACAAGCAAAT      50
AAATATGGTG TTCCAAGAAT AGTATTTGTA AACAAAATGG ATAGAATCGG      100
TGCAAAATTC TACAATGTAG AAGATCAAAT TCGCAACCGT TTAAAAGCTA      150
ATCCAGTTCC ACTTCAAATT CCAATCGGTG CTGAGGATAA TTTTAAAGGC      200
GTAATCGATC TTGTAACAT GAAAGCTTTA GTTTGGGAAG ATGATACTAA      250
GCCAACGGAT TATGTAGAAA AAGAAATTCC AGCTGAACCT AAAGAAAAGG      300
CAGAAGAATA TCGCACAAAA ATGATAGAAG CAGTTTCTGA AACTTCAGAT      350
GAGTTGATGG AAAAATATTT AGGCGGAGAA GAATTAAGCC TTGAAGAGAT      400
TAAAACAGGG ATTAAAGCAG GATGTTTAAG TCTTTCTATC GTTCCTATGC      450
TTTGCGGTAC AGCGTTTAAA AATAAAGGGG TTCAACCTTT ACTTGATGCT      500
GTTGTGGCTT ATTTACCAGC TCCTGATGAA GTGGCAAATA TCAAGGGTGA      550
ATATGAAGAT GGCACAGAAG TTTCTGTAAA ATCAACTGAT GATGGCGAGT      600
TTGCAGGACT TGCATTTAAA ATTATGACAG ATCCATTTGT AGGACAACCT      650
ACTTTCGTGC GTGTTTATCG CGGTTGTTTA GAAAGTGGTT CTTATGCTTA      700
TAACTCAACT AAAGATAAAA AAGAAAGAAT TGGTCGTTTG TTAAAAATGC      750
ACTCTAACCA AAGAGAAGAG ATTAAAGTGC TTTACGCAGG CGAAATTGGT      800
GCTGTTGTAG GACTTAAAGA TACTTTAACA GGGGATACTC TTGTAAGTGA      850
AAAAGATAAG GTAATCCTTG AAAGAATGGA TTTTCCAGAT CCAGTTATTT      900
CTGTTGCAGT TGAGCCAAAA ACTAAAGCAG ATCAAGAAAA AATGTCTATT      950
GCTTTAAATA AATTAGCACA AGAAGATCCA AGTTTtagag TTTCTACAGA     1000
TGAAGAAAGT GGCCAAACTA TCATTTcagg TATGGGTGAG TTACACCTTG     1050
AAATTATCGT TGATAGAATG CTTCTGTAAT TTAAAGTTGA AGCTGAAGTA     1100
GGTCAACCAC AAGTTGCTTA TCGCGAAACT ATTAGAAAAA CTGTTGAACA     1150
AGAATACAAA TACGCTAAAC AATCAGGTGG TCGTGgTCAG TATGGACATG     1200
TATTCTTACG CCTTGAACCA CTTGAGCCAG GTAGTGGATA TGAATTTGTT     1250
AATGATATCA AAGGTGGAGT AATTCCAAAA GAATACATTC CTGCAGTTGA     1300
TAAAGGTGTT CAAGAAGCAT TACAAAATGG TGTTTTAGCA GGTATCCTG     1350
TGGAAGATGT TAAAGTAACT GTTTATGATG GAAGTTATCA CGAGGTGGAT     1400
TCATCTGAGA TGGCGTTTAA ACTTGCTGCT TCTATGGGCT TTAAAGAAGG     1450
TGCTAGAAAA GCAGGTGCTG TGATCTTAGA GCCTATGATG AAAGTTGAAG     1500
TAGAAACTCC TGAAGATTAC ATGGGTGATG TTATTGGAGA TCTTAACAAA     1550
CGCCGTGGTC AAGTAAATAG CATGGATGAG CGTGGTGGAA ATAAAATCAT     1600
CACAGCATTT TG                                     1612

```

2) INFORMATION FOR SEQ ID NO: 2188

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1667 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Campylobacter jejuni* subsp. *jejuni*
 (B) STRAIN: ATCC 33560

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2188

```

ACAATTGAAG TTGAAAGATC TATGCGTGTT CTTGATGGTG CTGTTGCGGT      50
ATTTTGTTCa GTGGGTGGGG TGCAGCCTCA AAGTGAAACA GTTTGGAGAC      100
AAGCAAATAA ATATGGTGTT CCAAGAATAG TATTTGTAAA TAAAATGGAT      150
AGAATCGGTG CAAATTTCTA CAATGTAGAA GATCAAATTC GCAACCGTTT      200
AAAAGCTAAT CCAGTTCAC TTCAAATTC AATTGGTGCT GAGGATAATT      250
TTAAAGGCGT AATCGATCTT GTAACATGA AAGCTTTAGT TTGGGAAGAT      300
GATACTAAGC CAACGGATTA TGTAAGAAA GAAATTCCAG CTGAACCTAA      350
AGAAAAGGCA GAAGAATATC GCACAAAAAT GATAGAAGCA GTTTCTGAAA      400
CTTCAGATGA GTTGATGGAA AAATATTTAG GCGGAGAAGA ATTAAGCCTT      450
GAAGAGATTA AAACAGGGAT TAAAGCAGGA TGTTTAAGTC TTTCTATCGT      500
TCCTATGCTT TGCGGTACAG CGTTTAAAAA TAAAGGGGTT CAACCTTTGC      550
TTGATGCTGT TGTGGCTTAT TTACCAGCTC CTGATGAAGT GGCAAATATC      600
AAGGGTGAAT ATGAAGATGG CACAGAAGTT TCTGTAAAAT CAACTGATGA      650
TGGCGAGTTT GCAGGACTTG CATTTAAAAA CATGACAGAT CCATTTGTAG      700
GACAACTTAC TTTCTGCGT GTTTATCGCG GTTGTTTAGA AAGCGGTTCT      750
TATGCGTATA ACTCAACTAA AGATAAAAAA GAAAGAATTG GTCGTTTGTT      800
AAAAATGCAC TCTAACAAAA GAGAAGAGAT TAAAGTGCTT TACGCAGGCG      850
AAATTGGTGC TGTTGTAGGA CTTAAAGATA CTTTAACAGG GGATACTCTT      900
GCAAGTGAAA AAGATAAGGT AATCCTTGAA AGAATGGATT TTCCAGATCC      950
AGTTATTTCT GTTGCAGTTG AGCCAAAAAC TAAAGCTGAT CAAGAAAAAA      1000
TGTCTATTGC TTTAAATAAA TTAGCACAAG AAGATCCAAG TTTTAGAGTT      1050
TCTACAGATG AAGAAAGTGG TCAAACATATC ATTTTCAGGT TGGGTGAGTT      1100
ACACCTTGAA ATTATCGTTG ATAGAATGCT TCGTGAATTT AAAGTTGAAG      1150
CTGAAGTAGG TCAACCACAA GTTGCTTATC GCGAAACTAT TAGAAAAACT      1200
GTTGAACAAG AATACAAATA CGCTAAACAA TCAGGTGGTC GTGTCAGTA      1250
TGGACATGTA TTCTTACGCC TTGAACCACT TGAGCCAGGT AGTGGATATG      1300
AATTTGTTAA TGATATCAAA GGTGGGGTAA TTCCAAAAGA ATACATTCTT      1350
GCAGTTGATA AAGGTGTTCA AGAAGCATT AAAAAATGGTG TTTTAGCAGG      1400
TTATCCTGTG GAAGATGTTA AAGTAACTGT TTATGATGGA AGTTATCACG      1450
AGGTGGATTC ATCTGAGATG GCGTTTAAAC TTGCTGCTTC TATGGGCTTT      1500
AAAGAAGGTG CTAGAAAAGC AGGCGCTGTG ATCTTAGAGC CTATGATGAA      1550
AGTTGAAGTA GAAACTCCTG AAGATTATAT GGGTGATGTT ATTGGAGATC      1600
TTAACAAACG CCGTGGTCAA GTAAATAGCA TGGATGAGCG TGGTGAAAT      1650
AAAATCATCA CAGCATT      1667
  
```

2) INFORMATION FOR SEQ ID NO: 2189

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1255 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania guyanensis*

(B) STRAIN: ATCC 50126

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2189

| | | | | | |
|------------|------------|-------------|------------|-------------|------|
| TGACGGCGCT | GGACGTGACG | GAGGACCTCG | GCCGCGATGA | GCCGCTGACG | 50 |
| CTGGAGATCG | TGCAGCACTT | GGATGCGAAC | ACCGGCCGCT | GCATTGCGAT | 100 |
| GCAGACGACG | GACCTGCTGA | AGCTGAAGTC | GAAGGTTGTG | TCGACCGGCG | 150 |
| GCAACATCTC | CGTGCCGGTG | GGCCGCGAGA | CACTGGGCCG | CATCTTCAAC | 200 |
| GTGCTGGGCG | ACGCGATTGA | CCACCGCGGC | CCCGTGTGCG | AGAAGATGCG | 250 |
| CATGGCGATC | CACGCCGAGG | CGCCGAAGCT | GGCGGACCAG | GCTGCGGAGG | 300 |
| ACACGATCCT | GACGACCGGC | ATCAAGGTGA | TCGACCTGAT | TCTGCCCTAC | 350 |
| TGCAAGGGCG | GCAAGATCGG | CCTGTTTCGGC | GGTGCCGGTG | TGGGCAAGAC | 400 |
| TGTGATCATC | ATGGAGCTGA | TCAACAACGT | CGCGAAGGGG | CACGGCGGCT | 450 |
| TCTCCGTGTT | CGCCGGCGTT | GGCGAGCGCA | CGCGCGAGGG | CACGGACCTG | 500 |
| TACCTGGAGA | TGATGCAGTC | AAAGGTGATT | GACCTGAAGG | GCGAGTCGAA | 550 |
| GTGCGTGCTT | GTGTACGGGC | AGATGAACGA | GCCCCCGGGT | GCGCGCGCGC | 600 |
| GCGTTGCGCA | GTCTGCGCTG | ACGATGGCCG | AGTACTTCCG | CGACGTGGAG | 650 |
| GGCCAGAACG | TGCTGCTGTT | CATCGACAAC | ATCTTCCGCT | TCACGCAGGC | 700 |
| GAACTCCGAG | GTGTCTGCGC | TGCTGGGCCG | CATCCCGGCC | GCCGTGGGTT | 750 |
| ACCAGCCGAC | GCTTGCGGAG | GATCTTGCCA | TGCTGCAGGA | GCGCATTACG | 800 |
| TCGACGACGA | AGGGATCGAT | TACGTCTGTG | CAGGCTGTGT | ACGTGCCTGC | 850 |
| GGATGATATC | ACGGACCCCG | CGCCCGCGAC | GACGTTCTCG | CACCTGGACG | 900 |
| CGACGACTGT | GCTGGACCGC | GCGGTGGCGG | AGTCGGGCAT | CTACCCTGCC | 950 |
| GTGAACCCGC | TGGAGTGCGC | GTCGCGCATC | ATGGACCCCG | ATGTGATCGA | 1000 |
| CGTGGACCAT | TACAACGTTG | CACAGGATAT | CGTCCAGATG | CTGACCAAGT | 1050 |
| ACAAGGAGCT | GCAGGACATC | ATTGCGGTGC | TTGGCATCGA | CGAGCTGAGC | 1100 |
| GAGGAGGACA | AGGTCGTGGT | GGACCGCGCG | CGCAAGGTGA | CGCGGTTCCCT | 1150 |
| GTCGCAGCCG | TTCCAGGTTG | CGGAGGTGTT | CACCGGCATG | ACGGGCCACT | 1200 |
| ACGTGCAGCT | GAGCGACACG | GTGGAGTCGT | TCTCTGGCCT | GCTGATGGGG | 1250 |
| TCGTA | | | | | 1255 |

2) INFORMATION FOR SEQ ID NO: 2190

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1248 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma brucei* subsp. *brucei*
- (B) STRAIN: EATRO 795

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2190

| | | | | | |
|------------|-------------|------------|------------|------------|-----|
| GCCCTTGACG | TTGTTGACAA | ACTTGGCCGT | GATGAGCCTC | TTACTCTTGA | 50 |
| GATCGTGACG | CATCTTGACG | CCCACACGGG | CCGCTGTATC | GCGATGCAAA | 100 |
| CGACGGATCT | CCTCAAACCTG | AAGGCAAAGG | TCGTTTCGAC | AGGTGGCAAC | 150 |
| ATTTCCGTTT | CTGTCGGCCG | GGAAACACTA | GGTCGTATCT | TCAACGTGCT | 200 |
| TGGAGACGCT | ATTGACCAGC | GCGGCCCCGT | TGGTGAGAAA | CTGCGCATGC | 250 |
| CCATCCATGC | CGTGGCTCCC | AAGCTTGCGG | ACCAGGCCGC | TGAGGATGCG | 300 |
| GTGCTCACAA | CTGGTATTAA | GGTGATTGAT | CTCATTCTCC | CTTACTGCAA | 350 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| AGGTGGA | AAA | ATTGGCCTCT | TTGGGGGTGC | GGGTGTGGGC | AAAACCGTCA | 400 |
| TTATTATGGA | GCTCATTAAC | AACGTTGCCA | AGGGTCACGG | TGGTTTCTCT | | 450 |
| GTCTTCGCTG | GTGTTGGTGA | GCGTACCCGT | GAGGGAACGG | ATTTGTATCT | | 500 |
| TGAGATGATG | CAGTCTAAGG | TTATTGACCT | TAAGGGTGAG | TCCAAATGTG | | 550 |
| TGTTGGTGTA | CGGTCAGATG | AACGAGCCCC | CAGGTGCCCC | TGCGCGTGTT | | 600 |
| GCGCAGTCGG | CTCTGACGAT | GGCTGAGTAC | TTCCGTGATG | TGGAGGGCCA | | 650 |
| AGATGTGCTT | CTTTTATCG | ACAATATTTT | TCGTTTCACT | CAGGCTAACT | | 700 |
| CCGAGGTGTC | GGCGCTTCTG | GGTCGTATTC | CCGCCGCTGT | TGGCTACCAG | | 750 |
| CCTACCCTCG | CTGAGGATCT | AGGGCAGTTG | CAGGAGCGTA | TTACCTCAAC | | 800 |
| AACGAAAGGC | TCCATTACTT | CTGTGCAGGC | CGTATACGTG | CCGGCCGATG | | 850 |
| ACATTACCGA | TCCAGCTCCA | GCAACAACCT | TCTCACATCT | GGACGCCACA | | 900 |
| ACTGTGTTGG | ACCGTGCTGT | TGCCGAGTCT | GGTATCTACC | CCGCTGTTAA | | 950 |
| CCCACTGGAA | TGCGCCTCGC | GTATCATGGA | CCCCGACGTT | ATCAGTGTGG | | 1000 |
| ATCACTACAA | TGTTGCACAA | GATGTGGTAC | AGATGCTCAC | CAAGTACAGG | | 1050 |
| GAATTACAGG | ATATCATTGC | TGTCCTTGGT | ATCGACGAGC | TAAGCGAGGA | | 1100 |
| GGACAAACTT | ATCGTGGACC | GTGCGCGTAA | GTTGGTGAAG | TTCCTCTCCC | | 1150 |
| AGCCATTCCA | AGTTGCTGAG | GTCTTCACAG | GAATGACTGG | CCATTACGTG | | 1200 |
| CAGTTGGATG | ACACCATCGA | TTCCTTTTCT | GGTCTCCTCA | TGGGTACG | | 1248 |

2) INFORMATION FOR SEQ ID NO: 2191

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 440 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus nidulans*
- (B) STRAIN: WSA-176

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2191

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| GAGCTGATTG | TAAGTCATCG | ACCACTTGAT | ATATGAATAC | ATCTAACAGT | 50 |
| AGTAGAACAA | CATCGCCAAG | GCTCACGGTG | GTTACTCCGT | CTTCACTGGT | 100 |
| GTCGGTGAGC | GTACTCGTGA | GGGTAACGAT | CTGTACCACG | AAATGCAGGA | 150 |
| GACTGGTGTC | ATTGAGCTCG | ACGGCGAATC | CAAGGTGTCT | CTTGTGTTCTG | 200 |
| GTCAGATGAA | CGAGCCCCCA | GGTGCTCGTG | CCCGTGTCGC | CCTTACTGGT | 250 |
| CTGACCATCG | CCGAATACTT | CCGTGACGAG | GAGGGTCAGG | ACGTGCTGCT | 300 |
| CTTCATTGAC | AACATTTTCC | GTTTCACCCA | GGCCGGTTCC | GAGGTGTCTG | 350 |
| CCCTTCTTGG | TCGTATCCCC | TCTGCCGTCG | GTTACCAGCC | CACTCTGGCC | 400 |
| GTCGACATGG | GTGGTATGCA | GGAACGTATT | ACCACCACCA | | 440 |

2) INFORMATION FOR SEQ ID NO: 2192

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1262 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania panamensis*

(B) STRAIN: ATCC 50158

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2192

| | | | | | |
|------------|------------|------------|------------|------------|------|
| GCTGACGGCG | CTGGACGTGA | CGGAGGACCT | CGGCCGCGAT | GAGCCGCTGA | 50 |
| CGCTGGAGAT | CGTGCAGCAC | TTGGATGCGA | ACACCGGCCG | CTGCATTGCG | 100 |
| ATGCAGACGA | CGGACCTGCT | GAAGCTGAAG | TCGAAGGTTG | TGTCGACCGG | 150 |
| CGGCAACATC | TCCGTGCCGG | TGGGCCGCGA | GACGCTGGGC | CGCATCTTCA | 200 |
| ACGTGCTGGG | CGACGCGATT | GACCACCGCG | GCCCCGTGTG | CGAGAAGATG | 250 |
| CGCATGGCGA | TCCACGCCGA | GGCGCCGAAG | CTGGCGGACC | AGGCTGCGGA | 300 |
| GGACACGATC | CTGACGACCG | GCATCAAGGT | GATCGACCTG | ATTCTGCCCT | 350 |
| ACTGCAAGGG | CGGCAAGATC | GGCCTGTTCT | GCGGTGCCGG | TGTGGGCAAG | 400 |
| ACTGTGATCA | TCATGGAGCT | GATCAACAAC | GTCGCGAAGG | GGCACGGCGG | 450 |
| CTTCTCCGTG | TTCGCCGGCG | TTGGCGAGCG | CACGCGCGAG | GGCACGGACC | 500 |
| TGTACCTGGA | GATGATGCAG | TCAAAGGTGA | TTGACCTGAA | GGGCGAGTCG | 550 |
| AAGTGCGTGC | TTGTGTACGG | GCAGATGAAC | GAGCCCCCGG | GTGCGCGCGC | 600 |
| GCGCGTTGCG | CAGTCTGCGC | TGACGATGGC | CGAGTACTTC | CGCGACGTGG | 650 |
| AGGGCCAGAA | CGTGCTGCTG | TTCATCGACA | ACATCTTCCG | CTTCACGCAG | 700 |
| GCGAACTCCG | AGGTGTCTGC | GCTGCTGGGC | CGCATCCCCG | CCGCCGTGGG | 750 |
| TTACCAGCCG | ACGCTTGCGG | AGGATCTTGG | CATGCTGCAG | GAGCGCATTA | 800 |
| CGTCGACGAC | GAAGGGATCG | ATTACGTCTG | TGCAGGCTGT | GTACGTGCCT | 850 |
| GCGGATGATA | TCACGGACCC | CGCGCCCCCG | ACGACGTTCT | CGCACCTGGA | 900 |
| CGCGACGATA | GTGCTGGACC | GCGCGGTGGC | GGAGTCGGGC | ATCTACCCTG | 950 |
| CCGTGAACCC | GCTGGAGTGC | GCGTCGCGCA | TCATGGACCC | CGATGTGATC | 1000 |
| GACGTGGACC | ATTACAACGT | TGCACAGGAT | ATCGTCCAGA | TGCTGACCAA | 1050 |
| GTACAAGGAG | CTGCAGGACA | TCATTGCGGT | GCTTGGCATC | GACGAGCTGA | 1100 |
| GCGAGGAGGA | CAAGGTCGTG | GTGGACCGCG | CGCGCAAGGT | GACGCGGTTC | 1150 |
| CTGTGCGAGC | CGTTCCAGGT | TGCGGAGGTG | TTCACCGGCA | TGACGGGCCA | 1200 |
| CTACGTGCAG | CTGAGCGACA | CGGTGGAGTC | GTTCTCTGGC | CTGCTGATGG | 1250 |
| GGTCGTACGA | CC | | | | 1262 |

2) INFORMATION FOR SEQ ID NO: 2193

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 912 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus nidulans*

(B) STRAIN: WSA-176

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2193

| | | | | | |
|-------------|------------|------------|------------|------------|-----|
| TTCCGATGGT | CAAATGTACG | ATTGATATTC | CTTCCAGCCA | GTCAGGATAA | 50 |
| CAGCTGATAC | CAGTTGCAAA | TAGGCCCCAG | ACTCGTGAGC | ACTTGTGCT | 100 |
| TGCCCCGTCAG | GTTGGTGTCC | AGAAGATCGT | TGTCTTCGTC | AACAAGGTTG | 150 |
| ACGCTGTCTGA | TGACCCTGAG | ATGTTGGAGC | TTGTTGAGCT | CGAGATGCGT | 200 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GAGCTCCTCA | ACACTTACGG | TTTCGAGGGA | GAGGAGACCC | CTATCATCTT | 250 |
| CGGTTCCGCC | CTGTGCGCTC | TCGAAGGCCG | CCGCGAGGAC | ATTGGTACTC | 300 |
| AGCGTATTGA | CTCCCTCCTC | GAGGCCGTTG | ACACTTGGAT | CCCTACCCCC | 350 |
| CAGCGTGA | TGGACAAGCC | CTTCCTGATG | TCCATTGAGG | AAGTTTTCTC | 400 |
| CATTGGTGGT | CGTGGTACCG | TCGCCTCTGG | TCGTGTCGAG | CGTGGTCTCC | 450 |
| TCAAGAAGGA | TACCGAAGTT | GAAATTCACG | GTGCTGATGG | TATTCTGAAG | 500 |
| ACCAAGGTCA | CCGACATTGA | GACCTTCAAG | AAGAGCTGCG | ATGAGTCTCG | 550 |
| TGCTGGTGAC | AACTCCGGTC | TTCTCCTCCG | TGGTATCCGT | CGTGAGGATG | 600 |
| TTCGTCGTGG | TATGGTCATC | GCTGCCCTG | GCTCCATCAA | GGCCTCCAAG | 650 |
| AAGTTCATGG | TCTCCATGTA | CGTCTTGACT | GAGGCTGAAG | GTGGCCGCAA | 700 |
| GAACGGCTTC | GGTGCCAACT | ACCGCCCCCA | GGCTTTCATC | CGCACTGCTG | 750 |
| GTAAGTTTCG | AACTATTTGA | TTCATTGATC | ACGTCCCTAA | CTGTTACTTT | 800 |
| AGACGAGGCT | TGCGACCTTC | ATTTCCCTGA | TGAGGCCGAC | AAGGACCGCC | 850 |
| ACGTCATGCC | CGGTGACAAC | GTCGAAATGG | TCCTCAACCT | CAACAACCCC | 900 |
| GTTGCTGCTG | AG | | | | 912 |

2) INFORMATION FOR SEQ ID NO: 2194

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 887 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aureobasidium pullulans*
- (B) STRAIN: WSA-234

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2194

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TCATCGTCGT | CGCCGCTTCT | GACGGTCAGA | TGCCCCAGAC | CAGAGAGCAC | 50 |
| TTGCTTCTCG | CCCGTCAGGT | CGGTATCCAG | AAGATTGTCG | TCTTCGTCAA | 100 |
| CAAGGTCGAC | GCCATCGAGG | ACAAGGAGAT | GCTTGAGCTC | GTCGAGATGG | 150 |
| AGATGCGTGA | GCTCCTCAGC | AGCTACGGCT | TGAGGGGTGA | CGAGACCCCC | 200 |
| ATCGTCATGG | GTTCCGCTCT | TTGTGCCCTT | GAGAACCGCC | AGCCCGAGAT | 250 |
| TGGAACCACC | CAGATCGACA | ACCTGATGAA | CGCTGTCGAC | GAGTGGATCC | 300 |
| CCACTCCCCA | GAGAGATCTT | GAGAAGCCCT | TCCTCATGTC | CGTTGAGGAT | 350 |
| GTCTTCTCTA | TCCCCGGTCG | TGGTACTGTC | GTTTCTGGCC | GTGTTGAGCG | 400 |
| TGGTACCCTG | AAGAAGGATT | CCGAAATCGA | GCTTGTCGGC | AAGAACAAGG | 450 |
| TCCCCATCAA | GACCAAGGTC | ACCGACATCG | AGACCTTCAA | GAAGTCTTGC | 500 |
| GACGAGTCCC | GCGCTGGTGA | CAACTCCGGT | CTTCTGCTCC | GTGGTATCAA | 550 |
| GCGTGAGGAT | GTCAACCGTG | GTATGGTTGT | CGTCAAGCCC | GGTACCGTCA | 600 |
| CCTCGCACAA | GAAGTTCCTC | GTCTCCATGT | ACGTCCTGAC | CAAGGAGGAG | 650 |
| GGTGGTCGTC | ACACTGGTTT | CCACGGAAAC | TACCGTCCCC | AGATCTTCAT | 700 |
| CCGTACCGCT | GGTAAGTCCT | GACTTTGAAC | TGCTGACCAA | TTTTCGCATC | 750 |
| TCTAACATGT | TTTACAGACG | AGGCTGCCGC | TATTGACTGG | CCCGAAGGCA | 800 |
| CCGAGGACGC | TGACTCCAAG | ATGGTCATGC | CCGGTGACAA | CGTCGAGATG | 850 |
| GTCTGCTCTC | TTCACAGACC | TCTTGCCGTT | GAACAGG | | 887 |

2) INFORMATION FOR SEQ ID NO: 2195

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 984 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Emmonsia parva*
- (B) STRAIN: ATCC 10784

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2195

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TGACGGCCAA | ATGTAGGACT | CTTGCGCGGG | ACTGACTGAT | TGGGGGGAAT | 50 |
| CCATTTTTC | TTTTTCTTT | TTCTTTTGA | GTACATGATT | ATACTAATAT | 100 |
| TTGGATAACG | TACCAGGCCT | CAAACACGTG | AACATTTACT | CCTTGCCCGA | 150 |
| CAAGTCGGTG | TCCAGAAGAT | CGTCGTTTC | GTAAACAAGG | TCGATGTTCT | 200 |
| CGAAGATAAG | GAGATGTTGG | AGCTTGTCGA | GTTGGAAATG | AGAGAGCTCT | 250 |
| TGAACACCTA | CGGATTCGAG | GGTGAGGAAA | CGCCCATCAT | CTTCGGTTCT | 300 |
| GCCCTTTGTG | CCATGGAGGG | CCGCGAGCCT | GAGCTGGGCG | AGAAGAGAAT | 350 |
| TGATGAATTG | CTCGACGCTG | TTGATAGCTG | GATCCCTACG | CCGCAACGTG | 400 |
| ATACGGAAAA | ACCCTTCCTA | ATGTCCATTG | AGGAAGTGTT | CTCCATCTCC | 450 |
| GGTCGTGGAA | CCGTTGCCTC | CGGCCGTGTT | GAGCGTGGTG | TCCTCAAGAA | 500 |
| GGATTCCGAA | GTCGAGCTTG | TTGGCGGCGG | CGTCGCCCCA | ATCAGGACCA | 550 |
| AGGTAACCGA | TATCGAAACC | TTCAAGAAGT | CCTGCGACGA | GTCCAGGGCT | 600 |
| GGAGACAAC | CTGGCCTTTT | GTTGCGTGGT | GTCAAGCGTG | AGGATATCCG | 650 |
| CCGTGGTATG | GTCGTTGTCG | TTCCTGGCAG | CGTCAAGGCC | CACGACAAGT | 700 |
| TCTTGGTGTC | CATGTATGTT | CTGACCGAAG | CCGAGGGTGG | TCGCCGAAC | 750 |
| GGATTCCGCC | AAAACATATC | TCCTCAAATG | TTTATCCGCA | CAGCTGGTAC | 800 |
| GTAACGTATA | ATGCCTCTCC | TCTTCATATA | TACCACCCCC | CCCACCACTG | 850 |
| ACTCCCTGAC | TCTTCGATTA | CAGACGAGGC | CGCCGATCTC | AGCTTCCCTG | 900 |
| ACGCAGCAGA | CGAAACCAAA | CTGGTTATGC | CCGGTGACAA | CGTCGAGATG | 950 |
| ATCCTCAAGA | CACACCGCCC | CATAGCTGCC | GAAG | | 984 |

2) INFORMATION FOR SEQ ID NO: 2196

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 806 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Exserohilum rostratum*
- (B) STRAIN: WSA-215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2196

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GCCGCTCTG | ACGGCCAAAT | GCCCCAGACC | CGTGAGCACT | TGCTGCTTGC | 50 |
| TCGCCAGGTC | GGTGTTTCTG | AGATTGTTGT | TTTTGTCAAC | AAGGTCGATG | 100 |
| CCGTCGAGGA | CAAGGAGATG | TTGGAGCTCG | TTGAGATGGA | GATGCGTGAG | 150 |
| TTGCTCAGCA | GCTATGGCTT | CGAGGGCGAT | GAGACTCCCA | TTGTCATGGG | 200 |

| | | | | | |
|------------|------------|-------------|-------------|------------|-----|
| CTCCGCTCTC | TGCGCCATCG | AGGGCCGTGA | GCCCGAAATT | GGTGTCAACC | 250 |
| GTATTGATGA | GCTGCTCGAG | GCTGTCGACA | CCTGGATCCC | CACCCCTCAG | 300 |
| CGTGATACCG | ACAAGCCCTT | CCTCATGGCT | GTTGAGGACG | TCTTCTCCAT | 350 |
| TGCTGGCCGT | GGTACCGTCG | TTTCTGGCCG | TGTCGAGCGA | GGTGTCTTGA | 400 |
| AGCGCGATGC | CGAAGTCGAG | CTGGTTGGCA | AGGGCACTGC | GCCCATCAAG | 450 |
| ACCAAGGTTA | CCGACATTGA | AACCTTCAAG | AAGTCGTGTG | AAGAGTCTCG | 500 |
| CGCCGGTGAC | AACTCGGGTC | TCCTCCTCCG | TGGTGTCAAG | CGTGACGACG | 550 |
| TCCGCCGTGG | CATGGTCGTT | TCTGTTCTCTG | GACAAGTCAA | GGCCCACAAG | 600 |
| AAGTTCCTCG | TCTCCATGTA | TGTGTTGAGC | AAAGAGGAAG | GTGGCCGCCA | 650 |
| CACGGGCTTC | GGCGAGAACT | ACAGGCCACA | AATGTTTCATC | CGTACTGCTG | 700 |
| ACGAGTCATG | CGCGCTGCAC | TGGCCAGAAG | GTACCCCGA | TGCTCACGAC | 750 |
| AAGCTTGTTA | TGCCTGGTGA | TAACGTTGAG | ATGGTTTGTG | AGCTTCACGT | 800 |
| GCCACA | | | | | 806 |

2) INFORMATION FOR SEQ ID NO: 2197

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 799 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Fusarium moniliforme*
- (B) STRAIN: WSA-213

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2197

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ACCCGTGAAC | ACTTGCTCCT | CGCTCGTCAG | GTTGGTGTTC | AGCGAATTGT | 50 |
| CGTCTTTGTC | AACAAGGTCG | ATGCCATTGA | TGACCCCGAG | ATGCTTGAGC | 100 |
| TCGTCGAGAT | GGAGATGCGC | GAGCTTCTTA | ACACCTATGG | CTTCGAAGGC | 150 |
| GACGACACTC | CCGTCATCAT | GGGCTCGGCT | CTCATGTCTC | TCCAGAACCA | 200 |
| GCGCCCTGAG | ATTGGCACCG | AGAAGATTGA | TGAGCTTCTT | GCTGCCGTCG | 250 |
| ACGAGTGGAT | CCCAACCCCC | GAGCGTGACC | TTGACAAGCC | CTTCCTTATG | 300 |
| TCCGTCGAGG | ATGTCTTCTC | CATTGCTGGC | CGTGGTACCG | TCGTGTCTGG | 350 |
| CCGTGTGGAG | CGTGGTGTTC | TGAAGCGTGA | CCAGGAGATC | GAGCTTGTTG | 400 |
| GAAAGGGTCA | GGAGGTTATC | AAGACCAAGG | TTACCGACAT | CGAGACCTTC | 450 |
| AAGAAGTCTT | GTGAGCAGTC | CCAGGCTGGT | GACAACTCTG | GTCTCCTCAT | 500 |
| CCGAGGTGTT | CGCCGTGAGG | ATGTCCGCCG | TGGTATGGTC | GTCTGCGCTC | 550 |
| CTGGCACCGT | GAAGTCTCAC | ACCCAGTTCC | TCGCTTCCCT | CTATGTCCTC | 600 |
| TCCAAGGAGG | AGGGTGGCCG | ACACACCGGT | TTCCAGGAGC | ACTACCGACC | 650 |
| CCAGCTCTAC | CTCCGAACCG | CAGATGAGTC | CATTGACCTG | ACTTTCCTTG | 700 |
| AGGGTACTGA | GGATGCCTCC | AGCAAGATGG | TCATGCCTGG | CGACAACACC | 750 |
| GAGATGGTTG | TCACCATGGG | TCACCCCAAT | GCCATCGAGG | TTGGTCAGC | 799 |

2) INFORMATION FOR SEQ ID NO: 2198

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 819 bases
- (B) TYPE: Nucleic acid

- (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Fusarium solani*
(B) STRAIN: ATCC 32793

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2198

| | | | | | |
|------------|------------|-------------|------------|------------|-----|
| CTCTGACGGT | CAGATGCCCC | AGACCCGTGA | GCACTTGCTG | CTTGCCCGTC | 50 |
| AGGTCGGTGT | CCAGAAGATT | GTCGTCTTCG | TCAACAAGGT | CGATGCCATT | 100 |
| GACGACCCTG | AGATGCTTGA | GCTCGTCGAG | ATGGAGATGC | GTGAGCTCCT | 150 |
| CAACACCTAC | GGCTTCGAGG | GTGACGAGAC | CCCTGTCATC | ATGGGCTCTG | 200 |
| CTCTCATGTC | CCTCCAGAAC | CAGCGCCCCG | AGATCGGTAG | CCAGAAGATT | 250 |
| GACGAGCTCC | TTGCCGCCGT | TGACGAGTGG | ATCCCTACCC | CCGAGCGTGA | 300 |
| CCTTGACAAG | CCCTTCCTCA | TGTCCGTTGA | GGATGTCTTC | TCCATTGCCG | 350 |
| GCCGTGGTAC | CGTCGTCTCT | GGCCGTGTCTG | AGCGTGGTGT | CCTGAAGCGC | 400 |
| GACCAGGAGA | TTGAGCTCGT | CGGCAAGGGT | AACGAGGTCA | TCAAGACCAA | 450 |
| GGTCACCGAC | ATTGAGACCT | TCAAGAAGTC | TTGCGAGCAG | TCCCAGGCTG | 500 |
| GTGACAACTC | TGGTCTCCTC | ATCCGAGGTG | TCCGCCGTGA | GGATGTCCGC | 550 |
| CGTGGTATGG | TCGTCTGCGC | CCCCGGCACT | GTCAAGTCCC | ACACTCAGTT | 600 |
| CCTTTCTTCC | CTCTACGTCC | TCACCAAGGA | GGAGGGTGGC | CGACACACTG | 650 |
| GCTTCCAGGA | GCACTACCGA | CCCCAGCTCT | ACCTCCGAAC | TGCTGATGAG | 700 |
| TCCATCGACC | TGACCTTCCC | CGAGGGTACC | GAGGACGCCA | GCAGCAAGAT | 750 |
| GGTCATGCCC | GGTGACAACA | CCGAGATGGT | CATCACCATG | GGCCACCCCA | 800 |
| ACGCCATTGA | GGTCGGTCA | | | | 819 |

2) INFORMATION FOR SEQ ID NO: 2199

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1025 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Histoplasma capsulatum*
(B) STRAIN: WSA-377

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2199

| | | | | | |
|------------|------------|-------------|------------|------------|-----|
| TGCTGACGGC | CAAATGTAAG | ACGCCGCGAG | GGAGTGCTGA | AGGTTTTATG | 50 |
| CTTTTTAGGC | CCTTTGTTTC | TGAGAGCATG | ATGATACTAA | TACTCGGAAA | 100 |
| CGTATCTATT | AGGCCTCAAA | CACGTGAGCA | TTTGCTCCTT | GCCCGACAGG | 150 |
| TCGGTGTCCA | AAAGATCGTC | GTTTTTCGTGA | ACAAAGTCGA | CGCCCTTGAG | 200 |
| GACAAGGAGA | TGTTGGAGCT | TGTCGAGTTA | GAAATGAGAG | AGCTCTTAAA | 250 |
| CACCTACGGA | TTTGAGGGTG | AAGAGACACC | CATCATCTTT | GGTCTGCCCC | 300 |
| TTTGCGCCAT | GGAAGGCCGT | GAGCCTGAGT | TGGGAGAAAA | GAAAATTGAT | 350 |
| GAATTGCTGG | AGGCTGTTGA | TACTTGGATC | CCAACACCAC | AACGTGATAC | 400 |
| CGAAAAACCT | TTCTTGATGT | CCGTTGAGGA | AGTATTCTCT | ATCTCCGGTC | 450 |
| GTGGAACCGT | TGCCTCCGGT | CGTGTTGAGC | GCGGTGTCCT | CAAGAAGGAT | 500 |

| | | | | | |
|------------|------------|------------|------------|------------|------|
| TCAGAAGTCG | AGCTAATTGG | GGGCGGCTCC | ACCCCCATCA | GGACGAAGGT | 550 |
| AACTGATATC | GAAACTTTCA | AGAAATCCTG | TGACGAGTCT | AGAGCTGGGG | 600 |
| ACAAC TCCG | TCTTTTATTG | CGTGGTATCA | AGCGTGAAGA | TATCCGCCGT | 650 |
| GGTATGGTAG | TTGCCGTTCC | TGGCAGCGTC | AAGGCCCACG | ACAAGTTCTT | 700 |
| GGTGTGATG | TATGTCCTGA | CCGAAGCTGA | GGGTGGTCGC | CGAACCGGAT | 750 |
| TCGGCCAGAA | CTATCGTCCT | CAAATGTTCA | TCCGCACAGC | TGGTATGTCA | 800 |
| AAATGGGGCC | CCTTTTCATA | ATCCTTTCTT | TTTTTCCTTT | TCCTCTCTCT | 850 |
| ATCTCTCTCT | CTGTTTCTTT | TCAACTCGCC | TGATTACAGA | AATTAATAA | 900 |
| CCCCTTTGAT | TATAGACGAA | GCCGCCCATC | TCAGCTTCCC | TAGTGGAGCA | 950 |
| GATGAAAGCA | AACTCGTTAT | GCCTGGTGAC | AACGTCGAGA | TGATCCTCCA | 1000 |
| GACACACCGC | CCCGTGGCTG | CTGAG | | | 1025 |

2) INFORMATION FOR SEQ ID NO: 2200

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 667 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Kocuria kristinae*
- (B) STRAIN: ATCC 27570

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2200

| | | | | | |
|------------|------------|------------|-------------|------------|-----|
| AGCACGTGCT | GCTCGCCCGC | CAGGTCGGCG | TGCCGACCCT | GCTGGTCGCC | 50 |
| CTGAACAAGG | CCGACATGGT | CGAGGACGAG | GAGCTGCTGG | ACCTCGTCGA | 100 |
| GATGGAGGTC | CGCGAGCTGC | TGTCCTCCCA | GGAGTTCGAC | GGCGACAACG | 150 |
| CCCCCGTCAT | CCGCGTCTCC | GCGCTGAAGG | CGCTGGAGGG | CGACGAGAAG | 200 |
| TGGGTCAAGT | CCATCGAGGA | GCTCATGGAG | GCCGTGGACG | AGTACATCCC | 250 |
| GGACCCCGTG | CGCGACAAGG | ACAAGCCGTT | CCTGATGCCC | ATCGAGGACG | 300 |
| TCTTCACCAT | CACCGGGCGC | GGCACCCTGG | TGACCGGTCTG | CGCCGAGCGC | 350 |
| GGGACCCTGG | CCCTGAACTC | CGAGGTCGAG | ATCGTCGGCA | TCCGCCCGAT | 400 |
| CCAGAAGACC | ACGGTCACCG | GGATCGAGAT | GTTCCACAAG | CAGCTCGACG | 450 |
| AGGCCTGGGC | CGGCGAGAAC | TGCGGTCTGC | TGCTGCGCGG | CCTGAAGCGC | 500 |
| GACGACGTCG | AGCGCGGCCA | GGTCGTGGTG | AAGCCGGGTT | CCATCACCCC | 550 |
| GCACACCAAC | TTCGAGGCGA | ACGTCTACAT | CCTGTCCAAG | GACGAGGGTG | 600 |
| GGCGTCACAA | CCCGTTCTAC | TGAACTACC | GTCCGCAGTT | CTACTTCCGG | 650 |
| ACCACCGACG | TCACCGG | | | | 667 |

2) INFORMATION FOR SEQ ID NO: 2201

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 778 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Vibrio mimicus*
 (B) STRAIN: ATCC 33653

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2201

| | | | | | |
|-------------|------------|-------------|------------|------------|-----|
| TGCTGCAACA | GATGGTCCAA | TGCCACAAAC | TCGTGAGCAC | ATCCTGCTGG | 50 |
| GTCGCCAAGT | AGGTATTCCT | TACATCATCG | TATTCATGAA | CAAATGTGAC | 100 |
| ATGGTTGACG | ATGAAGAGCT | TCTAGAGCTG | GTTGAGATGG | AAGTTCGTGA | 150 |
| GCTTCTGTCT | GAGTACGATT | TCCCAGGTGA | TGACCTGCCA | GTAATCCAAG | 200 |
| GTTCAGCACT | AGGCGCGCTA | AACGGCGAAG | CACAGTGGGA | AGCGAAGATT | 250 |
| GTTGAACTAG | CAGAAGCACT | AGATTTCATAC | ATTCCAGAGC | CAGAGCGTGC | 300 |
| AGTAGACATG | GCATTCCTGA | TGCCAATCGA | AGACGTATTC | TCAATCCAAG | 350 |
| GTCGTGGTAC | AGTAGTAACT | GGCCGTATCG | AGCGCGGCAT | CCTGAAAGTG | 400 |
| GGTGACGAAG | TTGCGATCGT | TGGTATCAAA | GACACAGTAA | AAACTACCTG | 450 |
| TACAGGTGTA | GAAATGTTCC | GTAAGCTGCT | TGACGAAGGT | CGTGCAGGTG | 500 |
| AGAACGTTGG | TGCACTGCTA | CGTGGTACTA | AGCGTGAAGA | AGTAGAGCGT | 550 |
| GGTCAAGTAC | TGGCGAAGCC | AGGTTCAATC | ACCCACACAC | CTAAGTTCGA | 600 |
| ATCAGAAAGTA | TACGTACTGT | CAAAAGACGA | AGGTGGCCGT | CATACTCCAT | 650 |
| TCTTCAAAGG | TTACCGTCCA | CAGTTCTACT | TCCGTACAAC | TGACGTAAAC | 700 |
| GGCAGCATCG | AGCTTCCAGA | AGGCGTAGAA | ATGGTAATGC | CAGGCGACAA | 750 |
| CATCAAGATG | GTTGTAGACC | TGATTGCA | | | 778 |

2) INFORMATION FOR SEQ ID NO: 2202

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 412 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Citrobacter freundii*
 (B) STRAIN: ATCC 8090

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2202

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| AACGCTGACC | CTGCAGGTTA | TTGCCGCTGC | GCAGCGTGAA | GGTAAAACCT | 50 |
| GTGCATTTAT | CGATGCAGAA | CACGCACTGG | ACCCGGTCTA | TGCCCCGTAAG | 100 |
| CTTGGCGTTG | ATATCGATAA | CCTGCTGTGT | TCTCAGCCGG | ATACCGGTGA | 150 |
| ACAAGCGCTG | GAAATCTGTG | ATGCACTGGC | GCGCTCCGGT | GCGGTTGACG | 200 |
| TTATCGTTGT | CGACTCCGTT | GCCGCATTGA | CGCCGAAGGC | AGAAATCGAA | 250 |
| GGCGAGATTG | GCGACTCTCA | CATGGGCCTT | GCGGCGCGTA | TGATGAGCCA | 300 |
| GGCGATGCGT | AAGCTGGCCG | GTAACCTGAA | GCAGTCCAAC | ACGCTGCTGA | 350 |
| TTTTCATCAA | CCAGATCCGT | ATGAAGATTG | GCGTTATGTT | CGGTAACCCG | 400 |
| GAAACCACCA | CC | | | | 412 |

2) INFORMATION FOR SEQ ID NO: 2203

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 337 bases

- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium botulinum*
- (B) STRAIN: 20:1.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2203

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TTTAGATCCA | TCTTATGCTA | GAAATTTAGG | TGTTGATATA | GATAACCTAA | 50 |
| TAGTTTCTCA | ACCAGATACA | GGAGAACAGG | CTTTAGAGAT | AACAGAAGCT | 100 |
| TTAGTAAGAT | CAGGAGCAGT | AGATGTTATA | GTTGTAGACT | CTGTAGCAGC | 150 |
| TTTAGTTTCT | AGGGCAGAAA | TAGAAGGAGA | AATGGGAGAC | TCACATGTAG | 200 |
| GTCTTCAAGC | AAGACTTATG | TCTCAAGCCC | TAAGAAAATT | AGCAGGATCT | 250 |
| ATAAATAAAT | CTAAGTGTGT | AGCTATATTT | ATAAACCAAT | TAAGAGAAAA | 300 |
| GGTTGGTATA | ATGTTTGGAA | ATCCAGAAAC | AACTCCT | | 337 |

2) INFORMATION FOR SEQ ID NO: 2204

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 379 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Francisella tularensis*
- (B) STRAIN: ATCC 29684

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2204

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| AAAGCAAGGC | GGTACTGCAG | CATTTGTTGA | TGCTGAGCAT | GCACTAGATC | 50 |
| CAAAATATGC | AAAGCTTTTA | GGTGTTGATG | TTGATAATCT | GATCGTGTCA | 100 |
| CAGCCGGATA | CGGGTGAGCA | AGCTTTAGAG | ATTGCTGATA | TGTTGGTACG | 150 |
| TTCTGGAGGA | GTTGATATTG | TAGTAATTGA | CTCTGTTGCT | GCACTTACGC | 200 |
| CAAAGGCAGA | GATTGAGGGT | GACATGGGCG | ACTCGCACAT | GGGCTTACAA | 250 |
| GCAAGATTAA | TGTCACAAGC | ACTAAGAAAA | CTAACGGCAA | ATATCAAGCG | 300 |
| CTCAAATACT | CTAGTGATAT | TCATTAACCA | AATTCGTATG | AAGATCGGGG | 350 |
| TTATGTTTGG | TAACCCTGAA | ACTACAACT | | | 379 |

2) INFORMATION FOR SEQ ID NO: 2205

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 337 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Peptostreptococcus anaerobius*

(B) STRAIN: ATCC 27337

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2205

```

ACTTGACCCA GTATATGCAA GGGCTCTTGG AGTGGATATA GACAACCTAG-      50
TCATATCTCA GCCAGATACA GGAGAACAGG CCCTAGATAT AGCAGAGTCC-      100
CTTATAAGAT CAGGAGCTGT AGATATACTA GTAATAGACT CAGTAGCTGC-      150
CCTAGTACCT AAGGCAGAAA TAGAAGGTGA CATGGGAGAT TCTCACGTAG-      200
GTCTACAGGC TAGACTTATG TCACAGGCAC TTAGAAAATT GACTGGATCT-      250
ATAAAGAAGT CAAACTGTGT TGTATATTTT ATCAACCAGT TGAGAGAAAA-      300
AGTAGGGGTT ATGTTTCGGTA ATCCAGAGAC AACAAACA                      337

```

2) INFORMATION FOR SEQ ID NO: 2206

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 337 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Peptostreptococcus asaccharolyticus*

(B) STRAIN: LSPQ 2639

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2206

```

TCTTGATGCT GGATATGCAA AAAACCTTGG AGTAGATGTA GAAAATTTAA-      50
TTATTTCTCA ACCTGATACA GGTGAGCAAG CCTTAGAAAT AACTGAAGCT-      100
CTTGTAAGAT CTAACGCTGT TGATTTAATT ATTATAGACT CAGTTGCCGC-      150
ACTTGTACCA AAAGCAGAAA TCGATGGTGA CATGGGAGCT GCACAAATAG-      200
GTCTTCAAGC AAGACTTATG TCTCAAGCTC TTAGAAAATT AACTGGGGCA-      250
ATCAACAAGT CAAAATGTAC CGTTGTATTT ATTAACCAAC TTAGAGAAAA-      300
AGTTGGTATC ATGTTTGGTA ACCCAGAAAC TACAACA                      337

```

2) INFORMATION FOR SEQ ID NO: 2207

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 408 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Providencia stuartii*

1135

(B) STRAIN: ATCC 33672

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2207

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CTCACGTTGC | AAGTTATTGC | AGCAGCACAA | CGTAGCGGAA | AAACCTGTGC | 50 |
| ATTTATCGAC | GCTGAACATG | CGCTAGATCC | AATCTATGCG | AAAAAACTGG | 100 |
| GTGTTGATAT | CGATAACCTT | CTATGTTCTC | AACCTGATAC | TGGTGAGCAA | 150 |
| GCATTAGAGA | TTTGTGATGC | ACTGACGCGT | TCAGGCGCTG | TTGATGTCAT | 200 |
| TATCGTTGAC | TCCGTGGCCG | CATTAACACC | AAAAGCTGAA | ATTGAAGGTG | 250 |
| AAATCGGTGA | CTCACACATG | GGCTTAGCGG | CTCGTATGAT | GAGCCAAGCG | 300 |
| ATGCGTAAAT | TAGCGGGTAA | CTTAAAGAAC | TCGAATACAC | TTTTAATCTT | 350 |
| CATTAACCAA | ATCCGTATGA | AGATTGGCGT | TATGTTTGGT | AACCCAGAAA | 400 |
| CCACTACA | | | | | 408 |

2) INFORMATION FOR SEQ ID NO: 2208

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*
serotype Paratyphi A
- (B) STRAIN: ATCC 9150

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2208

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GCTGCAGGTG | ATTGCCGCTG | CGCAGCGTGA | AGGTAAAACC | TGTGCGTTTA | 50 |
| TCGATGCGGA | ACACGCGCTT | GACCCTGTTT | ACGCACGCAA | GCTGGGCGTC | 100 |
| GATATCGATA | ACCTGCTTTG | TTCTCAGCCG | GATACCGGCG | AGCAGGCGCT | 150 |
| GGAAATCTGT | GACGCGCTGG | CGCGTTCAGG | CGCGGTGGAC | GTCATTGTGG | 200 |
| TCGACTCCGT | AGCGGCGCTA | ACGCCGAAAG | CGGAAATCGA | AGGCGAAATT | 250 |
| GGCGACTCTC | ACATGGGCCT | CGCGGCGCGT | ATGATGAGCC | AGGCGATGCG | 300 |
| TAAGCTGGCG | GGGAACCTAA | AACAGTCCAA | CACGCTGTTG | ATTTTCATCA | 350 |
| ACCAGATCCG | TATGAAGATT | GGCGTGATGT | TCGGTAACCC | GGAAACCACC | 400 |
| ACC | | | | | 403 |

2) INFORMATION FOR SEQ ID NO: 2209

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 412 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*
serotype Typhimurium
(B) STRAIN: ATCC 14028

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2209

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GACCCTGACG | CTGCAGGTGA | TTGCCGCTGC | GCAGCGTGAA | GGTAAAACCT | 50 |
| GTGCGTTTAT | CGATGCGGAA | CACGCGCTTG | ACCCTGTTTA | CGCACGCAAG | 100 |
| CTGGGCGTCG | ATATCGATAA | CCTGCTCTGC | TCTCAGCCGG | ATACCGGCGA | 150 |
| GCAGGCGCTG | GAAATCTGTG | ACGCGCTGGC | GCGTTCAGGC | GCGGTGGACG | 200 |
| TCATTGTGGT | CGACTCCGTA | GCGGCGCTAA | CGCCGAAAGC | GGAAATCGAA | 250 |
| GGCGAAATCG | GCGACTCTCA | CATGGGCCTC | GCGGCGCGTA | TGATGAGCCA | 300 |
| GGCGATGCGT | AAGCTGGCGG | GGAACCTGAA | ACAGTCCAAC | ACGCTGTTGA | 350 |
| TTTTCATCAA | CCAGATCCGT | ATGAAGATTG | GCGTGATGTT | CGGTAACCCG | 400 |
| GAAACCACCA | CC | | | | 412 |

2) INFORMATION FOR SEQ ID NO: 2210

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 337 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus saprophyticus*
(B) STRAIN: ATCC 15305

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2210

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| ATTAGATCCT | GTTTACGCAG | AAGCTTTAGG | CGTAGATATA | CAAAATTTAT- | 50 |
| ATTTATCTCA | ACCTGATCAT | GGGGAACAAG | GTTTAGAAAT | TGCCGAAGCA- | 100 |
| TTTGTTAGAA | GTGGCGCTGT | TGATATCGTT | GTGGTCGATT | CAGTTGCTGC- | 150 |
| GCTTACACCT | AAAGCTGAAA | TTGAAGGTGA | AATGGGAGAT | ACGCACGTTG- | 200 |
| GTTTGCAAGC | ACGTCTTATG | TCCCAAGCCT | TGAGAAAGCT | TTCCGGTGCA- | 250 |
| ATTTCAAAAT | CAAATACAAC | AGCAGTATTT | ATCAACCAAA | TCCGTGAAAA- | 300 |
| AGTTGGTGTG | ATGTTCCGTA | ATCCTGAAGT | TACACCA | | 337 |

2) INFORMATION FOR SEQ ID NO: 2211

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 412 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Yersinia pseudotuberculosis*
(B) STRAIN: ATCC 29833

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2211

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GACACTGACA | TTACAGGTTA | TCGCCGCCGC | ACAGCGTGAA | GGCAAAACGT | 50 |
| GTGCATTTAT | CGATGCCGAA | CATGCCCTTG | ACCCAATCTA | TGCCAAGAAA | 100 |
| TTGGGTGTAG | ATATTGATAA | CCTACTGTGT | TCTCAGCCAG | ATACTGGCGA | 150 |
| GCAGGCACTG | GAAATTTGTG | ATGCGCTGAC | TCGCTCTGGT | GCGGTTGACG | 200 |
| TTATCATCGT | TGACTCCGTA | GCGGCATTGA | CACCAAAAGC | TGAAATTGAA | 250 |
| GGTGAAATTG | GCGATTCTCA | TATGGGCCTT | GCCGCGCGTA | TGATGAGCCA | 300 |
| GGCTATGCGT | AAGCTGGCGG | GTAACCTGAA | GAATGCGAAT | ACCTTACTGA | 350 |
| TTTTTATCAA | CCAAATCCGC | ATGAAAATTG | GCGTGATGTT | TGGTAACCCA | 400 |
| GAAACCACTA | CC | | | | 412 |

2) INFORMATION FOR SEQ ID NO: 2212

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 404 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zoogloea ramigera*
- (B) STRAIN: ATCC 25935

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2212

| | | | | | |
|------------|------------|-------------|-------------|------------|-----|
| CACGCTCACC | TTGCAAACCA | TCGCGGAAAT | GCAAAAACCTG | GGCGGCACCT | 50 |
| GCGCGTTTAT | CGACGCCGAG | CACGCACTGG | ACGTCACGTA | CGCGCAAAAG | 100 |
| CTGGGCGTCA | ACCTGAGCGA | TCTGCTGATC | TCGCAACCGG | ACACCGGCGA | 150 |
| ACAAGCGCTG | GAAATCTGCG | ACGCCCTGGT | GCGTTCCGGT | TCGGTGGACA | 200 |
| TGGTCGTGAT | CGACTCGGTC | GCCGCGCTGA | CCCCGCGCGC | CGAGATCGAA | 250 |
| GGCGACATGG | GCGATTCGCT | GCCAGGTTTG | CAGGCACGTT | TGATGTCGCA | 300 |
| AGCACTGCGC | AAGCTTACCG | GTTTCGATCAA | CCGCACCAAC | ACCCTGGTCA | 350 |
| TCTTCATCAA | CCAGATCCGC | ATGAAAATCG | GCGTCATGTT | CGGCAGCCCG | 400 |
| GAAA | | | | | 404 |

2) INFORMATION FOR SEQ ID NO: 2213

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2213

| | | | | |
|------------|------------|------------|------------|----|
| CGTGCCATTG | ACATGATTTC | CGAAGAAGAC | GCTGAAGGCA | CG |
|------------|------------|------------|------------|----|

42

2) INFORMATION FOR SEQ ID NO: 2214

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Abiotrophia adiacens*
- (B) STRAIN: ATCC 49175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2214

| | |
|--|-----|
| CAACTACATT ACGTTCTGCA ACACAAGGTC GTGGTACTTT CAGTATGACA | 50 |
| TTTGACCACT ATGAAGATGT TCCTAAGAGC ATTGCAGAAG AAATCATCAA | 100 |
| GAAAAATGGC GGTAACGGAG AATAA | 125 |

2) INFORMATION FOR SEQ ID NO: 2215

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Acinetobacter baumannii*
- (B) STRAIN: ATCC 19606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2215

| | |
|--|-----|
| CGACTCAAAT GCGTTCTATG TCTCAAGGTC GTGCGACATA CTCAATGGAA | 50 |
| TTTGCTAAAT ATGCTGAAAC TCCACGTAAC GTGGCTGAAG GCATCATCGC | 100 |
| TAAATTCCAA GCTGGCGGTA AAAAAGGTGA CGACGAGTAA | 140 |

2) INFORMATION FOR SEQ ID NO: 2216

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Actinomyces meyeri*
(B) STRAIN: ATCC 35568

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2216

| | |
|--|-----|
| CCGGTGACCT GCGTTCTAAG ACGCAGGGTC GCGCTGTCTA CTCCATGGAG | 50 |
| TTCGACAGCT ACGCCGAGGT TCCGCGCGCG GTCGCGGATG AGATCGTCGG | 100 |
| CAAGTCTCGG GGCAACTGA | 119 |

2) INFORMATION FOR SEQ ID NO: 2217

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium difficile*
(B) STRAIN: ATCC 9689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2217

| | |
|--|-----|
| CGACCGACGT TCGCTCCATG TCCCAGGGTC GCGCAAGCTA CTCTATGGAA | 50 |
| TTCAAAAAAT ACAACACAGC TCCGGCGCAC ATCGCTGAAA CTGTATCCAA | 100 |
| AAAACAAGGC TGA | 113 |

2) INFORMATION FOR SEQ ID NO: 2218

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Corynebacterium diphtheriae*
(B) STRAIN: ATCC 27010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2218

| | |
|--|-----|
| GCGACCTGCG TTCCCGTACC CAGGGCCGTG CAAACTACAC CATGATCTTC | 50 |
| GACTCCTACG CTGAGGTTCC TACCAACGTG GCAGCTGAGA TCGTGGCAGA | 100 |
| GCGCAACGGC ACTGCCTAA | 119 |

2) INFORMATION FOR SEQ ID NO: 2219

1140

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterobacter cloacae*
- (B) STRAIN: ATCC 13047

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2219

| | |
|--|-----|
| AACTCAGCTG CGTTCTCTGA CCAAAGGTCG TGCATCATAC ACCATGGAAT | 50 |
| TCCTGAAGTA TGATGATGCG CCTAACAACG TTGCTCAGGC CGTTATTGAA | 100 |
| GCCCGTGGTA AGTAA | 115 |

2) INFORMATION FOR SEQ ID NO: 2220

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella pneumoniae pneumoniae*
- (B) STRAIN: ATCC 13883

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2220

| | |
|--|-----|
| AACTCAGCTG CGTTCTCTGA CCAAAGGTCG TGCATCATAC ACCATGGAAT | 50 |
| TCCTGAAGTA TGATGATGCG CCGAACAACG TTGCTCAGGC CGTTATTGAA | 100 |
| GCCCGTGGTA AATAA | 115 |

2) INFORMATION FOR SEQ ID NO: 2221

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Listeria monocytogenes*
- (B) STRAIN: ATCC 15313

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2221

| | | | | | |
|------------|-------------|------------|------------|------------|-----|
| CAACTCACCT | TCG TTCAGGT | ACGCAAGGTC | GTGGTGTATA | CACTATGCAA | 50 |
| TTTGACCACT | ATGAAGAAGT | TCCTAAATCT | ATTGCTGAAG | AAATCATTAA | 100 |
| AGCTAATGGT | GGA | | | | 113 |

2) INFORMATION FOR SEQ ID NO: 2222

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium avium*
- (B) STRAIN: ATCC 25291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2222

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CGGCGACCTG | CGGTCCAAGA | CCCAAGGCCG | GGCGAACTAC | TCCATGGTCT | 50 |
| TCGACTCCTA | CGCCGAAGTG | CCGGCCAACG | TGTCGAAGGA | GATCATCGCG | 100 |
| AAGGCGACGG | GTCAGTGA | | | | 118 |

2) INFORMATION FOR SEQ ID NO: 2223

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium gordonae*
- (B) STRAIN: Mgor-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2223

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CCGGCGACCT | GCGGTCCAAG | ACGCAAGGCC | GGGCGAACTA | CTCCATGGTG | 50 |
| TTCGACTCGT | ACGCCGAAGT | TCCGGCGAAC | GTGTCCAAGG | AGATCATCGC | 100 |
| GAAGGCGACG | GGCGAATAG | | | | 119 |

2) INFORMATION FOR SEQ ID NO: 2224

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

1142

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium kansasii*

(B) STRAIN: Mkan-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2224

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CGGCGACCTG | CGGTCCAAGA | CTCAAGGCCG | GGCGAACTAC | TCGATGGTGT | 50 |
| TCGATTCTTA | CGCCGAAGTG | CCGGCTCAGG | TGTCGAAGGA | GATCATCGCG | 100 |
| AAGGCGACTG | GCGAGTGA | | | | 118 |

2) INFORMATION FOR SEQ ID NO: 2225

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 118 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium terrae*

(B) STRAIN: Mter-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2225

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CGGAGACTTG | CGGTCAAGA | CCCAGGGCCG | GGCGAACTAC | TCCATGGTGT | 50 |
| TCGACTCCTA | CGCCGAAGTG | CCGGCGCAGG | TGGCGAAGGA | GATTATCGCG | 100 |
| AAGGCAACGG | GCGAGTAA | | | | 118 |

2) INFORMATION FOR SEQ ID NO: 2226

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria polysaccharea*

(B) STRAIN: ATCC 43768

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2226

| | | | | | |
|------------|------------|-------------|------------|------------|-----|
| GACCGACCTG | CGTTCTGCAA | CCCAAGGCCG | CGCTACTTAC | TCTATGGAGT | 50 |
| TCAAGAAATA | TTCTGAAGCT | CCTGCCCCACA | TAGCTGCTGC | TGTAAGTAA | 100 |
| GCCCGTAAAG | GCTAA | | | | 115 |

2) INFORMATION FOR SEQ ID NO: 2227

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus epidermidis*
- (B) STRAIN: ATCC 14990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2227

| | |
|--|-----|
| AACTTCATTA CGTTCTAACA CGCAAGGTCG CGGTACTTAC ACAATGTACT | 50 |
| TTGACCACTA TGCAGAAGTT CCTAAATCAA TTGCTGAAGA AATCATCAAG | 100 |
| AAAAATAAAG GTGAATAA | 118 |

2) INFORMATION FOR SEQ ID NO: 2228

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus haemolyticus*
- (B) STRAIN: ATCC 29970

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2228

| | |
|--|-----|
| AACTTCATTA CGTTCTAACA CTCAAGGTCG CGGTACTTAC ACTATGTACT | 50 |
| TCGATCACTA TGCAGAAGTT CCAAATCAA TTGCTGATGA TATCATCAAA | 100 |
| AAAAATAAAG GTGAATAA | 118 |

2) INFORMATION FOR SEQ ID NO: 2229

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1630 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Succinivibrio dextrinosolvens*
 (B) STRAIN: ATCC 19716

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2229

```

CGTTCAATGC GTGTTCTAGA CGGTGCAGTT ATGGTTTACT GTGCTGTGGG      50
TGGTGTTCAG CCTCAGTCTG AAACCGTTTG GAGACAGGCT CAGAAGTACA      100
AGGTTCCCTCG TATTGCTTTC GTTAATAAGA TGGACCGTAC TGGTGCTAAT      150
TTCCTACGTG TTGTAGAGCA GATTAAGACC CGTCTAAAGG GTAACCCTGT      200
TCCTCTAATG TTACCTATCG GTAAAGAGGA CAGCTTTGTT GGTGTAGTTG      250
ACCTAATCAA GCGTAAGGCT ATCGACTGGG ATGAGGCATC TCAGGGTATG      300
AAGTTTGAGT ACGTTGACAT TCCAGCAGAT ATGGTTGAGG AAGTTGAAGA      350
GTGGCGTGCA AAGCTTGTAG AAGCAGCTGC AGAAGCTAAC GACGAGCTGA      400
TGGATAAATT CTTCCGTGGT GAAGAGCTGA CCGAGGAAGA GATCAAGGCT      450
GCTCTACGTG AGCGTACTCT TCGCAACGAA ATTATTCCTA TGTGCTGCGG      500
TTCAGCATTT AAGAACAAGG GTGTTCAGGC AATGCTTGAC GCTGTTGTTG      550
AGTATCTTCC ATCTCCAGCA GATGTTCCCTG CTGTTGAGGG TAAGACCCTA      600
ACCGGTGAAG CTGATACTCG TAAGGCTGAC GATAAAGAGC CATTCTCTGC      650
TTTAGCATTT AAGCTAGCAA ATGACCCATT CGTAGGTAAC TTAACATTCT      700
TACGTTGCTA CTCAGGCTTT ATTAAGTCTG GTGACACTGT AATGAACTCA      750
GATAAGCAGA AGCGTGAGCG TTTCGGCCGT CTAGTTCAGA TGCACGCTAA      800
TGCTCGTAAT GAGGTTAGCG AGGTTTATGC AGGTGACATC GTTGCTGCTA      850
TTGGTCTGAA GGAAACCGTT ACCGGTGATA CCTTATGTGA CCCAGAGCAT      900
CCAATCATTC TTGAGTCAAT CGACTTTGCA GAGCCAGTTA TCTCTGTAGC      950
AGTTGAGCCT AAGACCAAGG ACGATCAGGA GAAGATGGCT CTTGCTTTAC     1000
AGCGTTTAGC AAAAGAAGAT CCTTCATTCC GCGTTCGTAC AGACGAAGAG     1050
TCTGGCCAGA CCATTATTTT TGGTATGGGT GAGCTTCACC TAGACATCAT     1100
TGTTGACCGT CTACGCCGTG AGTTCAAGGT TGAGTGTAAT CAGGGTAAGC     1150
CACAGGTTGC ATACCGTGAG ACCATTAAGA GCAAGGTTGA ACAGCAAGGT     1200
AAGTTTGCTC GTCAGTCTGG TGGTCGTGGT CAGTACGGTG ACTGCTGGTT     1250
ACGTATGGAA CCTCTTGAGC CAGGTAAGGG CTACGAATTC GTGAATGAGA     1300
TTGTTGGTGG TGTAAATTCCT AAGGAATATA TCCCTGCAAT TGATAAGGGC     1350
TGTCAGGAGC AGATCGCTAA CGGTGTTCTA GCTGGTTTCC CAGTTGTTGA     1400
CATCAAGATC ACTGTATTCG ATGGTTCTTA CCACGAAGTT GACTCTTCAG     1450
AAATGGCATT CAAGATTGCT GCTTCTATGG CATTCAAAGA GGGCTTCAAG     1500
AAGGCAAATC CTGTTCTTCT AGAGCCTTTA ATGAAGGTAG AAGTTGATAC     1550
TCCTGAAGAC TACATGGGTG ACGTTATTGG TGAATTAAAC CGTCGTCGTG     1600
CTATCGTTGA AGGCATGGAA GATGGTCCTA

```

2) INFORMATION FOR SEQ ID NO: 2230

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1662 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Tetragenococcus halophilus*
 (B) STRAIN: ATCC 33315

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2230

| | | | | | |
|-------------|-------------|------------|------------|-------------|------|
| TTGAAGTGCA | ACGTTTCATTG | CGTGTGCTAG | ACGGTTCAGT | TACTGTCTTG | 50 |
| GACTCACAAT | CAGGTGTTGA | ACCACAAACT | GAAACAGTTT | GGCGTCAAGC | 100 |
| AACAGAATAT | CAAGTACCTC | GTATTGTATT | CTGTAATAAA | ATGGATAAGG | 150 |
| TTGGCGCAGA | TTTCTTATAT | TCTGTCCGAT | CATTGCATGA | TCGTTTAGAA | 200 |
| GCTAATGCAC | AACCTATCCA | ATTGCCAATT | GGTGCTGAAG | ATAACTTTGA | 250 |
| AGGTATCATC | GACCTTGTGA | AGATGAAAGC | TGAATTTTAT | AAAGATGATT | 300 |
| TAGGGACTAC | TTTCGAAGAA | ACTGAAATCC | CAGATGAATA | TAAAGAAACA | 350 |
| GCTCAAGAAT | GGCATAATAA | TTTGGTAGAA | TCTGTAGCTG | ATTTTGATGA | 400 |
| AGATATCATG | ATGAAATACT | TGGAAGGTGA | AGAAATTACA | CCTGAAGAGT | 450 |
| TACAAGCAGG | TATTCGTAAA | GCAACATTAT | CTGTTGAATT | TTACCCAGTA | 500 |
| TTATGTGGTT | CTGCATTTAA | AAACAAAGGT | GTTCAAATGA | TGTTGGATGC | 550 |
| AGTAATTGAT | TACTTGCCCT | CTCCAACCGA | CGTTCCCCCA | ATTAAAGGGA | 600 |
| TCGATCCGAA | AACAGATGAA | GAAACTGAAC | ATCCTGCTGA | TGATAGTGAG | 650 |
| CCTTTTTTCAT | CACTTGCTTT | TAAAGTTATG | TCAGACCCTT | ATGTTGGCCG | 700 |
| CTTAACTTTC | TTCCGTGTTT | ATTGAGGTGT | GTTGGATACA | GGTTCCTTATG | 750 |
| TATTGAATGC | TACTAAGGGT | TCACGTGAAC | GAATTGGTCG | TATTTTGCAA | 800 |
| ATGCATGCCA | ATTCTCGTTC | TGAGATCGAT | AAGGTTTATT | CAGGTGACAT | 850 |
| TGCAGCTGCT | GTAGGCTTGA | AGAACACTAC | AACAGGGGAT | ACCCTTTGTG | 900 |
| ATGAGAAAAA | TCCAGTTATT | TTGGAAACTA | TCAACTTCCC | TGAACCAGTA | 950 |
| ATTCAAGTTG | CTGTTGAACC | TAAGTCAAAA | GCTGACCAAG | ATAAAATGAG | 1000 |
| CGTAGCACTA | CAAAAACCTG | CAGAAGAAGA | CCCATCTTTT | AAAGTGGAAA | 1050 |
| CCAACGCTGA | AACTGGCGAA | ACTGTAATTG | CTGGTATGGG | TGAACTTCAA | 1100 |
| TTAGACGTTT | TTATTGACCG | TATGAAGACT | GAATTTAAAG | TGGATGCCAA | 1150 |
| TATTGGTGCA | CCACAAGTTT | CTTATCGTGA | AACTTTCCGT | TCATCAACTA | 1200 |
| AAGCTGAAGG | GAAATTTATC | CGCCAATCTG | GTGGTAGAGG | TCAATACGGT | 1250 |
| CACGTATGGG | TTGAATTTAC | TCCAACGAA | GAAGGAGCAG | GATTCGAATT | 1300 |
| TAAAAACTCC | ATTGTTGGTG | GGGTTGTCCC | TCGTGACTAT | ATACCTGCAG | 1350 |
| TACAAAAAGG | ACTGGAAGAC | GCCATGGAAA | ATGGTGTGTT | AGCTGGTTAT | 1400 |
| CCATTAGTTG | ACGTAAAGGC | AGAAGTGTTC | GATGGTTCTT | ACCATGACGT | 1450 |
| CGACTCTAAT | GAAACAGCCT | TCCGTATTGC | GGCTTCAATG | TCTCTACGTG | 1500 |
| AAGCTGCGAA | AAAGGCAGAT | CCAGTTATTC | TTGAACCGAT | GATGAAAGTA | 1550 |
| ACAATTAGTA | TCCCTGAAGA | ATATCTAGGT | GATATTATGG | GACATGTTAC | 1600 |
| AGCTCGTCGT | GGTCGTGTTG | AAGGAATGGA | TGCTCACGGT | AATGCACAAA | 1650 |
| CTGTAAATGC | GT | | | | 1662 |

2) INFORMATION FOR SEQ ID NO: 2231

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1652 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Veillonella parvula*
- (B) STRAIN: ATCC 10790

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2231

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| AAGTAGAACG | TTCTCTACGT | GTAAGTACG | GTTCTGTTGC | GGTGTTTCAGT | 50 |
| GCTAAAGGTG | GCGTTGAACC | TCAATCCGAA | ACAGTATGGC | GTCAGGCTTC | 100 |
| TAACTACGGC | GTACCTCGTA | TCGCTTATGT | AAATAAGATG | GATACTGTAG | 150 |

| | | | | | |
|------------|------------|-------------|------------|--------------|------|
| GTGCTGACTT | CTTCAACGTA | GTTGACATGA | TGAAAGCTCG | TTTGGGTGCA | 200 |
| AATTCCGTAG | CTATCCAAGT | ACCAATCGGT | GCTGAAGATA | CTTTCGAAGG | 250 |
| CATCATTGAC | TTGATGACTA | TGAAAGCGGA | AATTTATAAA | TCCGATGACG | 300 |
| GTAAAGAATA | TGAAATCACT | GATATCCCTG | CTGAATATCA | AGAAGTAGCA | 350 |
| GAAGCTCGTC | GCGAAATGAT | GATCGATGCT | ATCGCTGAAA | CAGATGATGA | 400 |
| TATCATGATG | AAATATTTGG | AAGGCGAAGA | AATTTCTGTG | GAAGAATTGA | 450 |
| AAGCGGCATT | GCGTAAAGCT | GTTATTGCTA | ACCAATTATT | CCCAGTTCTT | 500 |
| TGTGGTTCTT | CCTATAAAAA | TAAAGGTGTT | CAAATGTTAT | TGGATGCTGT | 550 |
| TATCGATTAC | ATGCCAGCTC | CAATCGACAT | CCCACCTATT | AAAGGTGTTG | 600 |
| TTCCTGGTAC | TGAAGAAGAA | ACAACCTCGTC | CTTCTTCCGA | TGAAGAGCCA | 650 |
| TTCTCTGCAT | TGGCATTCAA | AATCATGGCT | GACCCTTATG | TTGGTAAATT | 700 |
| AGCGTTCTTC | CGTGTGTACT | CCGGTACATT | GGAATCTGGC | TCCTACGTTT | 750 |
| TCAACTCCAC | TAAAGGTAAA | AAAGAACGTA | TCGGTCGTAT | TCTTCAAATG | 800 |
| CACGCTAACT | CCCGTAAAGA | AATCGAACGC | GTATATTCTG | GTGACATCGC | 850 |
| TGCGGCGGTT | GGCTTAAAGG | ATACTACTAC | AGGCGACACA | TTGTGTGATG | 900 |
| AAAAATCTCC | TGTAATCCTT | GAGTCCATGG | AATTCCCTGA | ACCAGTTATC | 950 |
| TCCGTTGCTG | TTGAACCTAA | AACAAAAGCT | GACCAAGAAA | AAATGGGTAC | 1000 |
| AGCTCTTGCT | CGTTTGGCAG | AAGAAGATCC | TACTTTCAAA | GTTTCGTA CTG | 1050 |
| ATGAAGAAAC | AGGTCAAAC | ATTATCTCTG | GTATGGGCGA | ACTTCACTTG | 1100 |
| GATATCATCG | TTGACCGTAT | GAACCGTGAA | TTCAAAGTAG | ATTGTAACGT | 1150 |
| AGGTAAACCT | CAAGTAGCAT | ACCGCGAAAC | TATCCGTAAA | GCTGTTAAGG | 1200 |
| CTGAAGGTAA | ATTTCGTACG | CAATCTGGTG | GTCGTGGTCA | ATATGGTCAC | 1250 |
| TGCTGGTTGG | AATTGATTCC | TCAAGAACCA | GGTGCTGGCT | TCGAGTTTGA | 1300 |
| AAACAAGGTT | TAGGTGGTG | CGATTCCCTG | TGAATACATC | GGACCTGTTG | 1350 |
| AAAGCGGTGT | TAAAGAAGCT | ATGGAATCCG | GTGTTATCGC | TGGGTACCCT | 1400 |
| ATGGTTGATG | TTAAAGTTAT | CGTATTTGAT | GGTTCTTACC | ATGACGTTGA | 1450 |
| CTCCAACGAA | ATGGCCTTCA | AAATTGCTGG | TTCTATGGGC | TTCAAAGAAG | 1500 |
| GTGCTCGCAA | AGCAGACCCT | GCATTGCTTG | AACCATATAT | GGCTGTAGAA | 1550 |
| GTAGACGTTT | CTGAAGAATA | CATGGGCGAC | GTTATCGGTG | ACTTGAAGTC | 1600 |
| TCGTCGTGGT | CGCATGGACG | GCATGGAAGC | TCGTAATGGT | TCCCAACATA | 1650 |
| TC | | | | | 1652 |

2) INFORMATION FOR SEQ ID NO: 2232

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1624 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Yersinia pseudotuberculosis*
- (B) STRAIN: ATCC 29833

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2232

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| AGTAGAGCGT | TCCATGCGTG | TTCTTGACGG | CGCGGTAATG | GTTTACTGTG | 50 |
| CAGTTGGTGG | TGTTCAGCCA | CAGTCTGAAA | CCGTATGGCG | TCAGGCTAAT | 100 |
| AAATATAAAG | TTCCACGTAT | TGCGTTCGTT | AACAAAATGG | ACCGTATGGG | 150 |
| TGCGAACTTC | CTGCGCGTAG | TTGGTCAACT | GAAATCTCGC | CTTGGTGCGA | 200 |
| ACCCAGTTCC | ACTGCAGTTG | GCAATTGGCG | CAGAAGAAAA | ATTCACCGGT | 250 |
| ATTATCGATC | TGGTGAAAAT | GAAAGCGATC | AACTGGAACG | AAGCTGATCA | 300 |
| GGGCGTGACC | TTCGAATATG | AAGAAATCCC | TGCTGATATG | GCTGAACTGG | 350 |

| | | | | | |
|------------|------------|------------|------------|------------|------|
| CTGCTGAATG | GCACCAGAAT | CTGGTTGAAT | CTGCGGCAGA | AGCGTCTGAC | 400 |
| GAGCTGATGG | ACAAATACTT | GGGTGGCGAA | GAGCTGACCG | AAGAAGAAAT | 450 |
| CAAGAAAGCT | TTACGTCAAC | GTGTTCTGAA | AAGCGAAATT | ATTCTTGTTA | 500 |
| CCTGTGGTTC | TGCGTTTAAA | AACAAAGGCG | TACAGGCAAT | GCTGGATGCG | 550 |
| GTTATTGAGT | ACCTGCCTGC | ACCAACTGAC | GTTGAATCAA | TCAACGGCAT | 600 |
| CTTGATGAT | GGCAAAGATA | CTCCGGCTGT | TCGTCATTCT | GACGACAAAG | 650 |
| AGCCGTTCTC | TGCTCTGGCG | TTCAAAATCG | CTACCGACCC | ATTCGTGGGT | 700 |
| AACCTGACGT | TCTTCCGCGT | GTACTCTGGT | ATTGTTAATT | CCGGTGATAC | 750 |
| CGTTCTGAAC | TCAGTGAAAT | CGCAACGTGA | ACGCTTAGGT | CGTATCGTAC | 800 |
| AGATGCACGC | TAACAAGCGT | GAAGAGATCA | AAGAAGTTCA | CGCCGGTGAT | 850 |
| ATCGCAGCCG | CTATCGGTCT | GAAAGATGTG | ACTACGGGTG | ACACTTTGTG | 900 |
| TGACCCGAAT | AATCCGATCA | TCTTGGAACG | TATGGAGTTC | CCAGAGCCGG | 950 |
| TAATCTCTGT | TGCTGTTGAA | CCAAAAACCA | AAGCTGACCA | AGAAAAAATG | 1000 |
| GGTATGGCTC | TGGGGCGTTT | GGCGAAAGAA | GATCCATCAT | TCCGCGTTTG | 1050 |
| GACTGACGAA | GAATCTGGTC | AGACTATCAT | CGCTGGTATG | GGTGAGTTGC | 1100 |
| ATTTGGATAT | CCTGGTTGAC | CGTATGCGCC | GCGAATTTAA | CGTGGAAGCA | 1150 |
| AACGTCGGTA | AACCTCAGGT | TGCGTACCGT | GAAACTATCC | GCGAAACCGT | 1200 |
| TAAGGATGTG | GAAGGTAAGC | ACGCTAAGCA | GTCAGGCGGT | CGTGGTCAGT | 1250 |
| ATGGTCATGT | TGTTATCGAC | ATGTCTCCAT | TGCCACCGGG | TGGTGTTGGG | 1300 |
| TATGAATTCG | TCAACGAAAT | CGTTGGTGGT | TCTATTCCTA | AAGAATTCAT | 1350 |
| TCCGGCCGTT | GATAAAGGTA | TTCAAGAACA | GCTGAAATCT | GGCCCTCTGG | 1400 |
| CAGGTTACCC | AGTTGTTGAC | GTTAAAGTGC | GTCTGCACTA | CGGTTCTTAC | 1450 |
| CATGACGTTG | ACTCCTCAGA | ATTGGCATT | AAATTAGCTG | GTTCTATCGC | 1500 |
| CTTTAAAGAA | GGTTTCAAAC | GAGCTAAACC | AGTTCTGCTT | GAGCCAATCA | 1550 |
| TGAAGGTTGA | AGTCGAAACC | CCTGAAGATT | ACATGGGTGA | CGTAATGGGC | 1600 |
| GACCTGAACC | GTCGTCGCGG | TATC | | | 1624 |

2) INFORMATION FOR SEQ ID NO: 2233

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1636 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Zoogloea ramigera*
- (B) STRAIN: ATCC 25935

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2233

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GAACGCTCGA | TGCGCGTGCT | TGACGGTGCT | TGCATGGTTT | ACTGCGCAGT | 50 |
| GGGCGGTGTT | CAGCCACAGT | CGGAAACCGT | GTGGCGTCAG | GCTAACAAGT | 100 |
| ACAAAGTGCC | ACGTCTGGCC | TTCGTCAACA | AGATGGACCG | TACCGGCGCC | 150 |
| AACTTCTTCA | AGGTGTACGA | GCAGATGCGT | GCTCGCCTGA | AGGCCAACCC | 200 |
| GGTCCTGATC | CAGATCCCTA | TCGGCGCTGA | AGACAACTTC | AAAGGCGTGA | 250 |
| TCGATCTGGT | CAAGATGAAG | GCTATCCTGT | GGGACGAAGC | GTGCAAGGC | 300 |
| ATGAAATTCG | ACTACGTCGA | TATTCCTGCA | GAGCTGGCTG | ATTCGGCCGC | 350 |
| CGAGTGCGCG | GAAAAGATGG | TTGAAGCTGC | TGCTGAAGCC | ACCGAAGAGC | 400 |
| TGATGAACAA | GTACCTGGAA | GAAGGCGACC | TGACCGAAGC | CGAGATCAAG | 450 |
| CAGGCGCTGC | GTACCCGTAC | CATCGCTTCG | GAAATCGTTC | CGATGATGTG | 500 |
| CGGTACCGCC | TTCAAGAACA | AGGGCGTACA | GGCCATGCTG | GACGCGGTCA | 550 |
| TCAATACCT | GCCATCGCCA | CTGGACATCG | ACGATGTCGG | CGGTACGGAC | 600 |

| | | | | | |
|------------|------------|------------|------------|------------|------|
| GAAGACGACC | AGCCAACCAC | CCGTCGCGCA | GCTGACGACG | AGAAATTCTC | 650 |
| GGCGCTGGCC | TTCAAGATCA | TGACCGACCC | GTTCGTCGGT | CAATTGGCCT | 700 |
| TCTTCCGCGT | GTACTCGGGC | GCCGTCAATT | CGGGCGACAC | CGTGTACAAC | 750 |
| TCGGTCAAAG | GTCGTAAAGA | GCGTCTGGGC | CGTATTCTGC | AGATGCACGC | 800 |
| GAATCAGCGC | GAAGAGATCA | AAGAAGTGCG | CGCCGGCGAC | ATCGCCGCTG | 850 |
| CGGTTGGCCT | GAAAGACGTG | ACCACGGGCG | AAACCCTGTG | CGATCCGACC | 900 |
| GCCATCATCA | CGCTGGAAAA | AATGATCTTC | CCTGAGCCTG | TGATTCAACA | 950 |
| GGCAGTCGAG | CCAAAAACCA | AGGCCGACCA | GGAAAAAATG | GGCCTGGCAC | 1000 |
| TGAACCGCCT | GGCACAGGAA | GATCCTTCGT | TCCGCGTGAA | GACCGATGAA | 1050 |
| GAATCGGGCC | AGACCATCAT | CGGTGGWATG | GGCGAGCTGC | ACCTGGAAAT | 1100 |
| TATCGTTGAC | CGCATGAAGC | GCGAATTCGG | CGTGGAAGCA | ACCGTCGGCA | 1150 |
| AGCCACAAGT | GGCTTACCGC | GAAACGATCC | GTAAAACCTG | CGAAGAATCG | 1200 |
| GAAGGCAAGT | TCGTCAAGCA | ATCCGGTGGT | CGTGGTCAAT | ACGGTCACGT | 1250 |
| TGTGCTGAAG | ATCGAGCCGC | AAGAACCAGG | CAAGGGCTTC | GAGTTCGTTG | 1300 |
| ACGCCATCAA | GGGCGGTACC | GTTCTTCGCG | AGTACATCCC | TGCGGTGGAA | 1350 |
| AAAGGCGTGC | GCGGCACCCT | GAACACCGGC | GTGCTGGCTG | GTTACCCGGT | 1400 |
| CGTGGACGTC | AAGGTCACGC | TGTTCTTCGG | TTCGTACCAC | GATGTGGACT | 1450 |
| CGAACGAAAA | CGCGTTCCAG | ATGGCCGCTT | CGATGGCATT | CAAAGAAGGC | 1500 |
| TGCCGCAAAG | CATCGCCAGT | CATTCTGGAG | CCAATGATGG | CTGTGGAAGT | 1550 |
| GGAAACGCCG | GAAGACTACG | CCGGTACCGT | GATGGGCGAC | CTGTCGTCCC | 1600 |
| GCCGCGGTAT | GGTGCAGGGC | ATGGACGAAA | TCCCAG | | 1636 |

2) INFORMATION FOR SEQ ID NO: 2234

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1656 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aeromonas hydrophila*
- (B) STRAIN: ATCC 7966

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2234

| | | | | | |
|------------|-------------|-------------|------------|-------------|-----|
| GGCACGAACG | ATCTTGCCGG | ACGGGCCATC | TTCCATACCT | TCGATCAGGC | 50 |
| CACGACGACG | GTTCAGGTCA | CCGATCACGT | CGCCCATATA | GTCTTCCGGA | 100 |
| GTCTCGACTT | CTACTTTCAT | GATCGGTTTCG | AGCAGAACCG | GGTTGGCCTT | 150 |
| CATGAAGCCA | GCCTTAAAGG | CCATGGAAGC | AGCGATCTTG | AACGCCAGTT | 200 |
| CGGAAGAGTC | GACATCGTGG | TAGGAACCGA | AGTGCAGACG | CACGCCCAGA | 250 |
| TCCATAACCG | GATAACCTGC | CAGCGGGCCG | GACTTCAGTT | GCTCGCGGAT | 300 |
| ACCCTTGTC | ACACCCGGGA | TGAACTCACC | AGGAATGACG | CCGCCCTTGA | 350 |
| TGTCGTTGAC | GAATTCGTAG | GCTTTGCCTT | CTTCCAGCGG | GTACATGTTCG | 400 |
| ATCACAACGT | GACCGTACTG | ACCACGACCA | CCGGACTGCT | TGGCGTGCTT | 450 |
| ACCTTCGATA | TCCTTGACGG | TGTTACGAAT | GGTTTCACGG | TAGGCAACCT | 500 |
| GCGGCTTACC | TACGTTTCGCT | TCTACCTTGA | ACTCGCGACG | CATACGGTCA | 550 |
| ACGATGATGT | CCAGGTGCAG | CTCACCCATA | CCGGCGATGA | TGGTTTGGCC | 600 |
| AGACTCTTCG | TCAGTCCATA | CGCGGAAGGA | CGGGTCTTCC | TGAGCCAGAC | 650 |
| GGCCCAGAGC | CAGGCCCATC | TTCTCTTGGT | CAGCCTTGGT | TTTCGGCTCA | 700 |
| ACTGCGATGG | AGATTACCGG | TTCCGGGAAT | TCCATACGCT | CGAGGATGAT | 750 |
| CGGCGCTTTT | TCGTCACACA | GGGTGTCACC | GGTGGTCACG | TCTTTCAGAC | 800 |
| CAATGGCGGC | AGCGATGTTCG | CCTGCGCGAA | CTTCTTTGAT | CTCTTCACGC | 850 |

| | | | | | |
|------------|------------|------------|-------------|-------------|------|
| TTGTTGGCGT | GCATCTGAAC | GATACGGCCA | AAACGCTCGC | GCTTCTCTTT | 900 |
| AACGGAGTTC | AGCACGGAGT | CACCGGAGTT | AACCACACCG | GAGTAAACGC | 950 |
| GGAAGAAGGT | CAGGTTGCCT | ACGAACGGGT | CGGTAGCAAT | CTTGAATGCC | 1000 |
| AGAGCAGCAA | ACGGCTCGTC | ATCAGAAGCA | TGACGCTCGT | CTTTGGTCTC | 1050 |
| GCCATCCAGC | TTCAGACCGT | CGATGGCTGC | TACGTCGGTC | GGCGCCGGCA | 1100 |
| GATAGTCAAC | CACGGCATCC | AGCATGGCCT | GTACGCCCTT | GTTCTTGAAC | 1150 |
| GCGGAGCCAC | AGGTAACCAG | GATGATTTTC | TTGTTTCAGAA | CACGCTGACG | 1200 |
| AAGAGCTTTC | TTGATCTCTT | CCTCGGTCAG | TTCTTCACCA | CCCAGGTATT | 1250 |
| TTTCCATCAG | GTCTTCAGAC | GCTTCAGCAG | CGGCTTCAAC | CAGGGTCATG | 1300 |
| CGCATTTCTT | GCGCTTTTTC | CAGCAGCTCA | GCCGGGACGT | CTTCGTAATC | 1350 |
| GAAGGATACG | CCCTGGTCAG | CTTCGCTCCA | GTTGATGGCT | TTTCATCTTGA | 1400 |
| CCAGGTCGAT | AACGCCCTTG | AAGTTCTCTT | CTGAACCGAT | GTTTCAGTTGC | 1450 |
| AGCGGAACCG | GGTTACCTTT | CAGACGGGTC | TTGATGTGCT | CAACGCAGCG | 1500 |
| CAGGAAGTTG | GCACCGGTAC | GGTCCATCTT | GTTGACGAAC | GCGATACGGG | 1550 |
| GAACCTTGTA | CTTGTTAGCC | TGACGCCATA | CGGTTTCAGA | CTGTGGCTGT | 1600 |
| ACGCCACHTA | CGGCACAGTA | CACCATCACG | GCGCCGTCCA | GAACACGCAT | 1650 |
| GGAACG | | | | | 1656 |

2) INFORMATION FOR SEQ ID NO: 2235

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Abiotrophia adiacens*
- (B) STRAIN: ATCC 49175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2235

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TCTCCTCCCC | ATTTGATAAC | TACCAAATGA | ACGCTATCGA | CTGGTTATGT | 50 |
| CAGTCATAAC | CAGTTGATTT | TTAAGAGAGT | TCTTTGGTAT | AATTACAATC | 100 |
| GGTAGATACT | GTTATAGAAT | CTAACAAAAC | TCAATTAATA | GGAGGAATCA | 150 |
| TTTAA | | | | | 155 |

2) INFORMATION FOR SEQ ID NO: 2236

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Acinetobacter baumannii*
- (B) STRAIN: ATCC 19606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2236

| | |
|---|----|
| TCTTTCGATT ACTATAAGCC CTAATAATT CTTAGTTAAA AACCAAGTGC | 50 |
| TCATGGAGCG ACCCTCCATG AGTAGTTAAT AAAGGAAGAT CATC | 94 |

2) INFORMATION FOR SEQ ID NO: 2237

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 150 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Actinomyces meyeri*
- (B) STRAIN: ATCC 35568

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2237

| | |
|---|-----|
| GTCCCACCGG GGGTGAGTGG GGGCCGGGCA ATCGGCCCCCT GCGAGCCGCC | 50 |
| TCACACACCA GTAATCCCAG TAGGTCTCAT GCCCATGGG TGTAAAGTG | 100 |
| ACACCTAGCC GTAGGCTGAG AATTCTACC CGAGTCCAGG AGGACGAAAA | 150 |

2) INFORMATION FOR SEQ ID NO: 2238

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium difficile*
- (B) STRAIN: ATCC 9689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2238

| | |
|----------------------------------|----|
| TTCAGTCCTT TAGGCAAGGA GTTAATTGTC | 30 |
|----------------------------------|----|

2) INFORMATION FOR SEQ ID NO: 2239

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 317 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Corynebacterium diphtheriae*

(B) STRAIN: ATCC 27010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2239

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CAACTAGCCT | CTCGCTTATC | TCCCTATGCA | AAATTTCTAG | TCAGTTACAG | 50 |
| GGCATAGGGG | GAGCGTAGGC | GGGGGTAGCG | GCTTGCTGAG | CACTTCCTCT | 100 |
| ACATCAAAGG | GAATGTTGAG | CCGGCCGTTA | CCCTGTACGA | TCCCATCTGG | 150 |
| TTTCTTCGGT | GGTTTGATAA | ATACCCCGTT | GTGACCCTAG | GATCATGTAA | 200 |
| CTGGCACAAT | GTAAATAGCT | GTACTGCCAG | GCTGCCGAAT | TAGCAGTCAG | 250 |
| AAATGTACAG | CACTGTCAAC | TCGTGGCTGC | GAAATCGTAG | CCACCACGAA | 300 |
| GTCCAGGAGG | ACACACA | | | | 317 |

2) INFORMATION FOR SEQ ID NO: 2240

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterobacter cloacae*

(B) STRAIN: ATCC 13047

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2240

| | | | | | |
|------------|------------|------------|------------|------------|----|
| TCCACAGGAT | TAAAACCTAA | GTCCCGTGCT | CTCTCCTGAG | GGGAGAGCAC | 50 |
| TATAGTAAGG | AATATAGCC | | | | 69 |

2) INFORMATION FOR SEQ ID NO: 2241

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Klebsiella pneumoniae* subsp. *pneumoniae*

(B) STRAIN: ATCC 13883

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2241

| | | | | | |
|------------|------------|------------|------------|------------|----|
| TCCACAGGAT | TAAAACCTAA | GTCCCGTGCT | CTCTCCGAAG | GGGAGAGCAC | 50 |
| TATAGTAAGG | AATATAGCC | | | | 69 |

2) INFORMATION FOR SEQ ID NO: 2242

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 126 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Listeria monocytogenes*
- (B) STRAIN: ATCC 15313

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2242

| | |
|--|-----|
| AACAACAAAG AAGATTAATT GATTTTTTTC GCAACATCAA GTATAACTTT | 50 |
| AGTTAGAAGT ATTACTTAGT TTAAATTTAA GCTAAGTAAA AAATAATTAT | 100 |
| CGAATTATCG AGGAGGATAT TTAAAA | 126 |

2) INFORMATION FOR SEQ ID NO: 2243

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium avium*
- (B) STRAIN: ATCC 25291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2243

| | |
|--|-----|
| GTGTCAACTC ACTGGCTCGG AGCCGAGCAA TCGGCTCAGC GAAGGCGACG | 50 |
| GGTCAGTAGC TACTGGCAGC GGAGTAATCT TGCCGGGTCA TTGGAATGCC | 100 |
| TTGGGCGCGG CACAACTGAA AACACCAACA CTGCTTTAAC AAGCACCAAC | 150 |
| TAGTCCAGGA GGACACAGAA | 170 |

2) INFORMATION FOR SEQ ID NO: 2244

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

1153

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium gordonae*
(B) STRAIN: Mgor-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2244

```
GGCCGGCAAG CCTGCGAGTA AGCTGACGCG GTTAGCACCG CGGCAAAACC      50
AAGAAAAATC AACACTGCTT TTTTAAGCAC CAACAGTCCA GGAGGACAAC      100
AAA                                                    103
```

2) INFORMATION FOR SEQ ID NO: 2245

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium kansasii*
(B) STRAIN: Mkan-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2245

```
GCCGAGCGTT GCGCGTAAGC TAGCTCGGTT ACCACGGCGG CAAAACCTAGA      50
AAAACATCAA CACTGCTTTT ATAAGCACCA ACAGTCCAGG AGGACACAGA      100
A                                                    101
```

2) INFORMATION FOR SEQ ID NO: 2246

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 91 bases
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium terrae*
(B) STRAIN: Mter-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2246

```
CCAGCTTCCC GCCCGTGCCG CTACGGTGGC ACAAACCCAA AAAGATCAAT      50
CCTGCTGTAA CCCAGCACCA ACAAGTCCAG GAGGACAAGA A                91
```

2) INFORMATION FOR SEQ ID NO: 2247

1154

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria polysaccharea*
- (B) STRAIN: ATCC 43768

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2247

| | |
|--|----|
| TCAGGCAAAT AGGCCGTCTG AAAGGCTGAA ATGATTTTTC AGACGGCATT | 50 |
| GTTCTTTAAT CGATCTTTAA TGTAAGGAA TTAGCTC | 87 |

2) INFORMATION FOR SEQ ID NO: 2248

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 218 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus epidermidis*
- (B) STRAIN: ATCC 14990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2248

| | |
|--|-----|
| TATAACTTGT TAAGACTAGC TATGCTAGGT TAAAATACAG GTTGAGCTTA | 50 |
| TTTATAAGCT GACATTTTAA TGATTTGATT TTTAGGGGTA AATGCATTAT | 100 |
| AAAAGAATTA TAAATTCTTT TATGCTACAC TCAATCAATT TTCTTCTCAT | 150 |
| GATGGTGAGA AACTATCATG AGAGATAAAT TTGAAATAAC TTTTATTAAG | 200 |
| AATAGGAGAG ATTTAATA | 218 |

2) INFORMATION FOR SEQ ID NO: 2249

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 204 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus haemolyticus*
- (B) STRAIN: ATCC 29970

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2249

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TTAAACTTGT | TTTAGCTAGA | ATTTCTAGGT | AAAATACAGC | GTAAGCTTAT | 50 |
| TAATTAAGCT | AACATCTTTA | TGAATTGATT | TTTTACTGAA | AATGCATTAT | 100 |
| AAATGAATTA | TGAATTCTAA | CAATCATTAT | GTCTCATGAT | GGTGAGAAAC | 150 |
| TATCATGAGA | GATAATATTG | AAATAACTTT | TACTAGAATA | GGAGAGATTT | 200 |
| AATA | | | | | 204 |

2) INFORMATION FOR SEQ ID NO: 2250

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2250

CCGTCGGATG GTGTCGTATA CCGCGGAGTC GCCGACGG 38

2) INFORMATION FOR SEQ ID NO: 2251

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2251

CGGAGCCGTT CTCGCTGCGT TACATGCTGG TGGCTCCG 38

2) INFORMATION FOR SEQ ID NO: 2252

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2252

CCACATACAG TGTCTCTC 18

2) INFORMATION FOR SEQ ID NO: 2253

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2253

CATTACCCAA CCGAAAGTA

19

2) INFORMATION FOR SEQ ID NO: 2254

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2254

ACCTGAACAG AGAGAAATG

19

2) INFORMATION FOR SEQ ID NO: 2255

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Abiotrophia adiacens*
- (B) STRAIN: ATCC 49175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2255

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGGCAAAAG | AAAAATTTGA | CCGTTCAAAA | CCACACGTTA | ACATTGGTAC | 50 |
| AATCGGCCAC | GTTGACCACG | GTAAAACAAC | ATTAAGTGCT | GCTATCACAA | 100 |
| CTGTTTTAGC | TAAGAAAGGT | TTCGCGCAAG | CTCAAGATTA | CGGTTCAATC | 150 |
| GATAAAGCTC | CAGAAGAACG | CGAACGTGGT | ATCACAATCA | ACACTTCTCA | 200 |
| CGTTGAGTAC | GAAACAGACA | CTCGTCACTA | TGCTCACGTT | GACTGCCCAG | 250 |
| GACACGCGGA | CTACGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2256

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Acinetobacter baumannii*
- (B) STRAIN: ATCC 19606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2256

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGGCTAAAG | CCAAGTTTGA | ACGTAATAAA | CCACACGTAA | ACGTGGGTAC | 50 |
| AATCGGTCAC | GTTGACCATG | GTAAAACAAC | TTTAACTGCT | GCGATTGCAA | 100 |
| CAATTTGTGC | AAAAACTTAC | GGCGGTGAAG | CGAAAGATTA | CTCACAAATC | 150 |
| GACTCAGCAC | CTGAAGAAAA | AGCAGCTGGT | ATTACAATTA | ATACATCACA | 200 |
| CGTAGAATAC | GATTCTCCAA | CTCGTCACTA | CGCACACGTT | GACTGCCCAG | 250 |
| GCCACGCCGA | CTACGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2257

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Actinomyces meyeri*
- (B) STRAIN: ATCC 35568

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2257

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGGCGAAGG | CCAAGTTTGA | GCGCACCAAG | CCGCACGTCA | ACATCGGCAC | 50 |
| GATTGGTCAC | GTTGACCACG | GCAAGACGAC | GCTGACGGCA | GCTATCACCA | 100 |
| AGGTGCTGCA | TGACAAGTAC | CCCGAACTGA | ACGAGTTCAC | CCCCTTCGAT | 150 |
| CAGGTCGACA | ACGCTCCCGA | GGAGCGCGAT | CGTGGCATCA | CGATCAACGT | 200 |
| CTCTCACGTT | GAGTACCAGA | CCGAGGCGCG | TCACTACGCG | CACGTTGACG | 250 |
| CTCCCGGCCA | CGCCGACTAC | GTCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2258

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases

- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium difficile*
- (B) STRAIN: ATCC 9689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2258

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGGCTAAAG | AAAAATTTGA | TCGTTCCCTA | CCGCACGTCA | ACGTTGGCAC | 50 |
| TATCGGTCAC | GTTGACCATG | GTAAAACCAC | TCTGACTGCT | GCTCTGACTC | 100 |
| GCGTTTGCTC | CGAAGTATTC | GGTTCCGCAA | TCGTTGATTT | CGATAAAATC | 150 |
| GACAGCGCAC | CAGAAGAAAA | AGCTCGTGGT | ATCACCATCA | ACACCGCGCA | 200 |
| CGTTGAATAC | AACTCGCTGA | TCCGTCAC | CGCTCACGTT | GACTGCCCAG | 250 |
| GTCACGCTGA | CTATGTGAAG | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2259

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Corynebacterium diphtheriae*
- (B) STRAIN: ATCC 27010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2259

| | | | | | |
|-------------|------------|------------|------------|------------|-----|
| GTGGCAAAGG | CTAAGTTCGA | GCGTACCAAG | CCGCACGTCA | ACATCGGCAC | 50 |
| CATCGGTCAC | GTTGACCACG | GTAAGACCAC | CACCACCGCT | GCTATCACCA | 100 |
| AGGTTTTTGGC | AGACGCTTAC | CCAGAGCTGA | ACGAAGCTTT | CGCTTTCGAT | 150 |
| GCCATCGATA | AGGCACCGGA | AGAGAAAGAG | CGTGGTATTA | CCATCAACAT | 200 |
| CTCCACCGTG | GAGTACCAGA | CCGAGAAGCG | CCACTACGCA | CACGTTGACG | 250 |
| CTCCAGGTCA | CGCTGACTAC | ATCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2260

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterobacter cloacae*
 (B) STRAIN: ATCC 13047

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2260

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGTCTAAAG | AAAAATTGGA | ACGTACAAAA | CCGCACGTCA | ACGTTGGTAC | 50 |
| TATCGGCCAC | GTTGACCATG | GTAAAACTAC | CCTGACTGCT | GCAATCACTA | 100 |
| CCGTTCTGGC | TAAAACCTAC | GGTGGTTCTG | CTCGTGCATT | CGACCAGATC | 150 |
| GATAACGCAC | CAGAAGAAAA | AGCTCGTGGT | ATCACCATCA | ACACCTCTCA | 200 |
| CGTTGAATAT | GACACCCCGA | CTCGCCACTA | CGCACACGTA | GACTGCCCAG | 250 |
| GTCACGCCGA | CTATGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2261

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Klebsiella pneumoniae* subsp. *pneumoniae*
 (B) STRAIN: ATCC 13883

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2261

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGTCTAAAG | AAAAATTGGA | ACGTACAAAA | CCGCACGTCA | ACGTTGGTAC | 50 |
| TATCGGCCAC | GTTGACCATG | GTAAAACTAC | CCTGACTGCT | GCCATCACTA | 100 |
| CCGTTCTGGC | TAAAACCTAC | GGTGGTTCCG | CTCGCGCATT | CGACCAGATC | 150 |
| GATAACGCAC | CGGAAGAAAA | AGCTCGTGGT | ATCACCATCA | ACACCTCTCA | 200 |
| CGTTGAATAT | GACACCCCGA | CTCGCCACTA | CGCGCACGTA | GACTGCCCCG | 250 |
| GCCACGCCGA | CTATGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2262

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Listeria monocytogenes*
 (B) STRAIN: ATCC 15313

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2262

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGGCAAAAG | AAAAATTGGA | CCGCTCTAAA | CCCCATGTTA | ACATTGGTAC | 50 |
| TATTGGACAC | GTTGACCATG | GTAAACAAC | TTTAACTGCT | GCAATTACAA | 100 |

| | | | | | |
|------------|------------|-------------|------------|------------|-----|
| CTGTACTTGC | TAAAAAAGGC | TATGCTGATG | CACAAGCTTA | TGACCAAATT | 150 |
| GATGGTGCTC | CAGAAGAAAG | AGAACGTGGA | ATCACAATCT | CTACTGCTCA | 200 |
| CGTTGAGTAC | CAAAGTGACA | GCCGTCACCTA | TGCACACGTT | GACTGCCCAG | 250 |
| GACATGCCGA | TTACGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2263

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium avium*
- (B) STRAIN: ATCC 25291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2263

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGGCGAAGG | CGAAGTTCGA | GCGGACGAAG | CCGCACGTCA | ACATCGGGAC | 50 |
| CATCGGTCAC | GTTGACCACG | GCAAGACCAC | GCTGACCGCG | GCTATCACCA | 100 |
| AGGTTCTGCA | CGACAAGTAC | CCGGACCTGA | ACGAGTCCCG | CGCGTTCGAC | 150 |
| CAGATCGACA | ACGCGCCCGA | GGAGCGTCAG | CGCGGTATCA | CCATCAACAT | 200 |
| CTCCACGTG | GAGTACCAGA | CCGACAAGCG | GCACTACGCT | CACGTCGACG | 250 |
| CCCCGGGTCA | CGCCGACTAC | ATCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2264

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium gordonae*
- (B) STRAIN: Mgor-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2264

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGGCGAAGG | CGAAGTTCCA | GCGGACCAAG | CCGCACGTCA | ACATCGGGAC | 50 |
| CATCGGTCAC | GTTGACCACG | GCAAGACCAC | TCTGACCGCG | GCTATCACCA | 100 |
| AGGTCCTGCA | CGACAAGTAC | CCGGACCTGA | ACGAGTCCAA | GGCGTTCGAC | 150 |
| CAGATCGACA | ACGCGCCTGA | GGAGCGTCAG | CGCGGTATCA | CGATCAACAT | 200 |
| CGCGCACGTG | GAATACCAGA | CCGAGAAGCG | TCACTACGCG | CACGTCGACG | 250 |
| CCCCCGGCCA | CGCCGACTAC | ATCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2265

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium kansasii*
- (B) STRAIN: Mkan-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2265

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GTGGCGAAGG | CGAAGTTCCA | GCGGACCAAG | CCCCACGTCA | ACATCGGGAC | 50 |
| CATCGGTCAC | GTTGACCACG | GCAAGACCAC | CCTGACCGCG | GCTATCACCA | 100 |
| AGGTCCTGCA | TGACAAGTTC | CCGGACCTGA | ACGAGTCGAA | GGCGTTCGAC | 150 |
| CAGATCGACA | ACGCTCCTGA | GGAGCGCCAG | CGCGGTATCA | CGATCAACAT | 200 |
| CGCGCACGTG | GAGTACCAGA | CCGAGAAGCG | GCACTATGCA | CACGTCGACG | 250 |
| CGCCGGGCCA | CGCCGACTAC | ATCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2266

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium terrae*
- (B) STRAIN: Mter-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2266

| | | | | | |
|------------|-------------|------------|------------|------------|-----|
| GTGGCGAAGG | CGAAGTTCTGA | GCGGACGAAG | CCGCACGTCA | ACATCGGGAC | 50 |
| CATCGGTCAC | GTTGACCACG | GCAAGACCAC | GCTGACCGCG | GCTATCACCA | 100 |
| AGGTTCTGCA | CGACAAGTAC | CCGGACCTCA | ACGAGTCGCG | TGCGTTCGAC | 150 |
| CAGATCGACA | ACGCTCCCGA | AGAGCGTCAG | CGCGGTATCA | CCATCAACAT | 200 |
| CTCCACGTG | GAGTACCAGA | CCGAGAAGCG | GCACTACGCC | CACGTCGACG | 250 |
| CTCCTGGTCA | CGCTGACTAC | ATCAAGAAC | | | 279 |

2) INFORMATION FOR SEQ ID NO: 2267

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

1162

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria polysaccharea*
 (B) STRAIN: ATCC 43768

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2267

```

ATGGCTAAGG AAAAATTCGA ACGTAGCAAA CCGCACGTAA ACGTTGGCAC      50
CATCGGTCAC GTTGACCATG GTAAAACCAC TCTGACTGCT GCTTTGACTA      100
CTATTTTGGC TAAAAAATTC GGCGGTGCTG CAAAAGCTTA CGACCAAATC      150
GACAACGCAC CCGAAGAAAA AGCACGCGGT ATTACCATTA ACACCTCGCA      200
CGTAGAATAC GAAACCGAAA CCCGCCACTA CGCACACGTA GACTGCCCCG      250
GTCACGCCGA CTACGTTAAA AAC                                     273

```

2) INFORMATION FOR SEQ ID NO: 2268

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus epidermidis*
 (B) STRAIN: ATCC 14990

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2268

```

ATGGCAAAAG AAAAATTTGA TCGCTCAAAA GAACATGCCA ATATTGGTAC      50
TATCGGTCAC GTTGACCATG GTAAAACAAC TTAAACAGCT GCTATCGCAA      100
CTGTATTAGC TAAAAATGGT GACACTGTTG CACAATCATA CGATATGATT      150
GACAACGCTC CAGAAGAAAA AGAACGTGGT ATTACAATCA ATACTGCACA      200
TATCGAATAC CAAACTGACA AACGTCATA TGCTCACGTT GACTGCCCCAG      250
GACACGCTGA CTATGTTAAA AAC                                     273

```

2) INFORMATION FOR SEQ ID NO: 2269

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 bases
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: Double
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus haemolyticus*
 (B) STRAIN: ATCC 29970

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2269

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGGCAAAAG | AAAAATTTGA | TCGCTCAAAA | GAACATGCCA | ATATTGGTAC | 50 |
| TATCGGTCAC | GTTGACCATG | GTAAACTAC | TTTAACAGCT | GCTATCGCAA | 100 |
| CTGTATTAGC | TAAAAATGGT | GACACTGTAG | CACAATCATA | TGACATGATT | 150 |
| GACAACGCTC | CAGAAGAAAA | AGAACGTGGT | ATCACAATCA | ATACTGCACA | 200 |
| CATCGAGTAT | CAAACTGACA | AACGTCACTA | TGCTCACGTT | GACTGCCCAG | 250 |
| GACACGCTGA | CTATGTTAAA | AAC | | | 273 |

2) INFORMATION FOR SEQ ID NO: 2270

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 812 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aeromonas hydrophila*
- (B) STRAIN: ATCC 7966

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2270

| | | | | | |
|------------|-------------|------------|------------|-------------|-----|
| CGATCCTGGT | AGTAGCAGCG | ACTGACGGCC | CGATGCCGCA | GACTCGTGAG | 50 |
| CACATCCTGC | TGGGTCGTCA | GGTAGGCGTT | CCGTACATCA | TCGTGTTTCAT | 100 |
| GAACAAGTGT | GACATGGTAG | ATGACGAAGA | GCTGCTGGAA | CTGGTCGAGA | 150 |
| TGGAAGTTCG | CGAACTGCTG | TCCGAGTACG | ACTTCCCGGG | TGATGACCTG | 200 |
| CCGGTAGTCC | GTGGTTCYGC | ACTGAAAGCG | CTGGAAGGCG | AAGCTCAGTG | 250 |
| GGAAGAGAAG | ATCCTGGAAC | TGGCTGGCCA | CCTGGACACC | TACATTCCGG | 300 |
| AGCCGGAGCG | TGCCATCGAC | CTGCCGTTCC | TGATGCCTAT | CGAAGACGTA | 350 |
| TTCTCCATCG | CTGGCCGYGG | TACCGTAGTG | ACCGGTCGTG | TAGAGCGCGG | 400 |
| TATCGTCAAA | GTTGGTGAAG | AAGTGGAAAT | CGTKGGTATC | AAAGATACCA | 450 |
| CCAAGACCAC | CTGTACCGGC | GTTGAAATGT | TCCGCAAAC | GCTGGACGAA | 500 |
| GGTCGTGCAG | GCGAGAACAT | CGGTGCACTG | CTGCGTGGCG | TGAAGCGTGA | 550 |
| AGACGTAGAG | CGTGGTCAGG | TACTGGCCAA | GCCGGGCACC | ATCAAGCCGC | 600 |
| ACACCAAGTT | YGAATCTGAA | GTGTACGTGC | TGTCCAAAGA | AGAAGGTGGT | 650 |
| CGTCATACCC | CGTTCCTTCAA | AGGCTACCGT | CCGCAGTTCT | ACTTCCGTAC | 700 |
| TACCGACGTG | ACCGGTACCA | TCGAACTGCC | GGAAGGCGTA | GAGATGGTAA | 750 |
| TGCCGGGCGA | CAACATCAAG | ATGGTTGTTA | CCCTGATTGC | GCCGATCGCG | 800 |
| ATGGACGACG | GC | | | | 812 |

2) INFORMATION FOR SEQ ID NO: 2271

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 799 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bilophila wadsworthia*

(B) STRAIN: ATCC 49260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2271

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CGACGGTCCC | ATGCCCCAGA | CCCGTGAGCA | CATCCTGCTC | GCCCGTCAGG | 50 |
| TCGGCGTGCC | TCACCTCGTC | GTGTTCATGA | ACAAGTGTGA | CCTCGTCGAC | 100 |
| GACCCCGAAC | TGCTCGAACT | CGTCGAAATG | GAAGTCCGCG | AACTGCTGAG | 150 |
| CTCCTACGGC | TACCCCGGCG | ATGAAATCCC | GGTTGTCCGC | GGTTCCGCTC | 200 |
| TGAAGGCTCT | GGAATCCGAT | AGCGCTGATT | CCCCTGACGC | CCAGTGCGTG | 250 |
| CTCGAACTGC | TCGCCGCTTG | CGACAGCTAC | TTCCCGGATC | CGGTCCGCGA | 300 |
| AACCGACAAG | CCCTTCCTGA | TGCCCATCGA | AGACGTGTTC | TCCATCTCCG | 350 |
| GCCGCGGTAC | CGTGGTCACC | GGTCGTGTGG | AACGTGGCAT | CATCAAGGTC | 400 |
| GGCGAAGAAG | TCGAAATCGT | GGGTATCCGT | CCCACCGTGA | AGACGACCTG | 450 |
| CACCGGCGTC | GAAATGTTCC | GCAAGCTGCT | CGATCAGGGC | CAGGCCGGCG | 500 |
| ACAACATCGG | CGTCTGCTC | CGCGGCACGA | AGCGTGACGA | AGTGGAACGC | 550 |
| GGCCAGGTTT | TCGCCGCTCC | CAAGAGCATC | ACGCCCCACA | AGAAGTTCAA | 600 |
| GGCTGAAGTG | TACGTTCTGT | CCAAGGAAGA | AGGCGGCCGC | CATACCCCGT | 650 |
| TCTTCACCGG | CTATCGTCCT | CAGTTCTACT | TCCGTACCAC | CGACATCACC | 700 |
| GGTATCATCG | CTCTTGAAGA | AGGCGTTGAA | ATGGTTATGC | CCGGCGATAA | 750 |
| CGTACCTTT | AATGTCGAGC | TCATTACCC | CATCGCCATG | GAAAAGGGC | 799 |

2) INFORMATION FOR SEQ ID NO: 2272

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 786 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Brevundimonas diminuta*
- (B) STRAIN: ATCC 11568

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2272

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| GATCCTGGTG | TGCTCGGCCG | CTGACGGCCC | GATGCCGCAG | ACCCGCGAGC | 50 |
| ACATCCTGCT | GTCGCGTCAG | GTCGGCGTTC | CGGCGCTGGT | GGTGTTCTTG | 100 |
| AACAAGGTCG | ACATGGTCGA | CGACGAGGAR | CTGCTGGAGC | TGGTCGAGAT | 150 |
| GGAAGTGCGC | GAGCTGCTGA | GCTCGTACCA | GTTCCCGGGC | GACGACATCC | 200 |
| CGGTGGTCAA | GGGCTCGGCC | CTGGCCGCGG | TGGAGGATCG | CGACCCGCAC | 250 |
| ATCGGCGCCG | AGCGCGTTCT | GGAGCTGATG | GCGGCGGTCT | ACAGCTACAT | 300 |
| CCCGCAGCCG | GAACGTCCGA | TCGACATGCC | GTTCTTGATG | CCGGTGGAAG | 350 |
| ACGTGTTCTC | GATCTCGGGC | CGCGGCACCG | TGGTGACGGG | TCGCGTCGAG | 400 |
| CGCGGCGTCG | TCAAGGTCGG | TGAAGAAGTC | GAAATCGTCG | GCATCCGTCC | 450 |
| GGTTCAGAAG | ACGACCTGCA | CGGGCGTCGA | AATGTTCCGC | AAGCTGCTGG | 500 |
| AYCAGGGTCA | RGCCGGCGAC | AACGTGGGCG | TGCTGCTGCG | CGGCACCAAG | 550 |
| CGTGAAGACG | TCGAGCGCGG | CCAGGTGCTG | TGCAAGCCGG | GTTTCGATCAC | 600 |
| CCCGCACACC | AAGTTCGTGG | CTGAAGCCTA | CATCCTGAAC | AAGGAAGAAG | 650 |
| GCGGCCGTCA | CACGCCGTTC | TTCACGAACT | ACCGTCCGCA | GTTCTACTTC | 700 |
| CGCACGACGG | ACGTGACCGG | CATCGTGCGC | CTGAAGGAAG | GCGTCGAGAT | 750 |
| GATCATGCCG | GGCGACAACG | CCGAGCTGGA | CGTCGA | | 786 |

2) INFORMATION FOR SEQ ID NO: 2273

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 560 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*
- (B) STRAIN: LSPQ 2583

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2273

| | | | | | |
|------------|-------------|------------|------------|-------------|-----|
| GCTATTATGG | CTGGATTACT | ATCCAATACG | CTATCCAAGA | ATCCCGTAAC | 50 |
| GTACCTGCCG | TCAAATCGCT | GGAAGCAGTC | GGATTAGATA | ATTCATTGAA | 100 |
| ATTCCTCAAT | GGCCTTG GTA | TCAATTATCC | TGAGATGCAT | TATTCTAATG | 150 |
| CGATTTCAAG | TAATACAAGC | GAATCTGGTA | ACCAATACGG | AGCAAGTAGC | 200 |
| GAAAAAATGG | CTGCCGCTTA | CGCTGCCTTT | GCTAATGGCG | GTACATATTA | 250 |
| CAAACCGCAA | TACGTCAACC | GAGTTGTCTT | TAGCGACGGT | ACAGAAAAAG | 300 |
| TCTTTTCAAA | TGGCGGATCA | AAAGCCATGA | AAGAGACGAC | AGCCTACATG | 350 |
| ATGACAGACA | TGATGAAGAC | CGTTCTTCAG | TCTGGAAGTG | GTACCAATGC | 400 |
| TGCAATCCCA | GGAGTTTATC | AAGCAGGTAA | AACTGGTACT | TCCAACATATG | 450 |
| CAGATGATGA | GCTAGAGAAG | TTGACAAAAC | CTTATTACAG | TTCTAGCATT | 500 |
| GTCACACCAG | ACGAACTATT | TGTTGGCTAT | ACTCCACAAT | ACTCTATGGC | 550 |
| TGTTTGGACA | | | | | 560 |

2) INFORMATION FOR SEQ ID NO: 2274

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*
- (B) STRAIN: ATCC 49456

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2274

| | | | | | |
|------------|-------------|------------|------------|------------|-----|
| GCTACTTTGG | AAACATTACT | GTCCAATATG | CTCTTCAACA | ATCACGTAAT | 50 |
| GTCACAGCCG | TTGAAACTTT | GAATAAGGTC | GGTCTAGATA | AGGCTAAAGC | 100 |
| CTTCCTTAAT | GGGCTTG GTA | TTGATTATCC | AAGCATGCAT | TATGCAAACG | 150 |
| CCATTTCAAG | TAATACAAC T | GAATCCAACA | AAAAATACGG | TGCAAGTAGT | 200 |
| GAAAAAATGG | CTGCTGCCTA | CGCTGCTTTT | GCTAATGGTG | GTATTTACCA | 250 |
| CAAGCCAATG | TACATCAATA | AAATCGTCTT | TAGCGACGGT | AGTGAGAAAG | 300 |
| AATTTTCTGA | TGCCGGCACA | CGAGCTATGA | AAGAACTAC | TGCCTATATG | 350 |
| ATGACTGAAA | TGATGAAAAC | AGTCCTAGTA | TACGGTACCG | GACGTGGAGC | 400 |
| CTACCTACCA | TGGCTTCCAC | AAGCAGGTAA | GACAGGTACT | TCTAACTATA | 450 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CTGACGACGA | AATTGAAAAG | TATATCAAGA | ACACTGGCTA | CGTAGCCCCA | 500 |
| GATGAAATGT | TTGTAGGGTA | TACTCGTAA | TATGCAATGG | CTGTTTGGAC | 550 |
| A | | | | | 551 |

2) INFORMATION FOR SEQ ID NO: 2275

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 560 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*
- (B) STRAIN: ATCC 903

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2275

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| GCTATTATGG | CTGGATTACT | ATCCAATACG | CTATCCAAGA | ATCCCGTAAC | 50 |
| GTACCAGCCG | TCAAATCGCT | GGAAGCAGTC | GGATTAGATA | ATTCATTGAA | 100 |
| GTTCTTCAAT | GGCCTTGGTA | TTAATTACCC | TGAAATGCAT | TATTCTAATG | 150 |
| CGATTTCAAG | TAATACAAGC | GAATCTGGTA | ACCAATACGG | AGCAAGTAGC | 200 |
| GAAAAAATGG | CTGCCGCTTA | CGCTGCCTTT | GCTAATGGCG | GTACATATTA | 250 |
| CAAACCGCAA | TACGTCAACC | GAGTTGTCTT | TAGCGACGGT | ACAGAAAAAG | 300 |
| TCTTTTCAAA | TGGCGGATCA | AAAGCCATGA | AAGAAACGAC | AGCCTACATG | 350 |
| ATGACAGACA | TGATGAAGAC | CGTTCTTCAA | TCTGGAAGTG | GTACCAATGC | 400 |
| TGCAATTCCA | GGAGTCTATC | AAGCAGGTAA | AACCGGCACT | TCCAACTATG | 450 |
| CAGATGATGA | ACTAGAGAAG | TTGACAAAAC | CTTATTACAG | TTCTAGCATT | 500 |
| GTCACACCAG | ACGAGCTGTT | TGTTGGCTAC | ACTCCACAGT | ACTCTATGGC | 550 |
| TGTTTGGACA | | | | | 560 |

2) INFORMATION FOR SEQ ID NO: 2276

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 550 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus oralis*
- (B) STRAIN: ATCC 35037

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2276

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| TTATTTTGGG | AATATCACCA | TCCAATATGC | GCTCCAACAA | TCACGGAACG | 50 |
| TTACAGCCGT | AGAAACCTTG | AACAAAGTCG | GTTTGGATAG | AGCCAAGACC | 100 |
| TTCCTGAATG | GAATCGGTAT | TGACTATCCA | GATATGCACT | ATGCCAACGC | 150 |
| GATTTCAAGT | AATACGACTG | AGTCAAACAA | AAAGTACGGA | GCAAGTAGTG | 200 |

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| AGAAAATGGC | TGCTGCTTAC | GCTGCTTTTG | CTAACGGTGG | TATCTACCAT | 250 |
| AAACCAATGT | ATATCAACAA | AATCGTCTTT | AGCGATGGTA | GCTCAAAAGA | 300 |
| ATACGCTGAT | CCTGGTACTC | GTGCCATGAA | AGAGACGACC | GCCTATATGA | 350 |
| TGACAGAAAT | GATGAAGACT | GTCTTGGCAT | ACGGAACGGG | TCGTGGTGCT | 400 |
| TATCTCCCTT | GGCTACCTCA | AGCTGGTAAG | ACTGGTACAT | CAAACCTATAC | 450 |
| AGATGATGAA | ATTGAAAAC | ACATCAAAAA | TACTGGTTAT | GTAGCCCCAG | 500 |
| ACGAAATGTT | TGTTGGTTAT | ACTCGCAAAT | ATTCAATGGC | TGTWTGGACA | 550 |

2) INFORMATION FOR SEQ ID NO: 2277

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 356 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 35401

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2277

| | | | | | |
|------------|------------|------------|------------|-------------|-----|
| GCTCCTATCT | GGATTATGCG | ATGTCGGTCA | TTGTTGGCCG | TGCGCTGCCA | 50 |
| GATGTCCGAG | ATGGCCTGAA | GCCGGTACAC | CGTCGCGTAC | TTTACGCCAT | 100 |
| GAACGTACTA | GGCAATGACT | GGAACAAAGC | CTATAAAAAA | TCTGCCCCGTG | 150 |
| TCGTTGGTGA | CGTAATCGGT | AAATACCATC | CCCATGGTGA | CTCGGCGGTC | 200 |
| TATGACACGA | TCGTCCGCAT | GGCGCAGCCA | TTCTCGCTGC | GTTATATGCT | 250 |
| GGTAGACGGT | CAGGGTAACT | TCGGTTCTAT | CGACGGCGAC | TCTGCGGCGG | 300 |
| CAATGCGTTA | TACGGAAATC | CGTCTGGCGA | AAATTGCCCA | TGAACTGATG | 350 |
| GCCGAT | | | | | 356 |

2) INFORMATION FOR SEQ ID NO: 2278

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 23511

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2278

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| CTCTTATCTG | GATTATGCGA | TGTCGGTCAT | TGTTGGCCGC | GCGCTGCCGG | 50 |
| ATGTCCGAGA | TGGCCTGAAG | CCGGTACACC | GTCGCGTACT | TTACGCCATG | 100 |
| AACGTATTGG | GCAATGACTG | GAACAAAGCC | TACAAAAAAT | CAGCCCGTGT | 150 |
| CGTTGGTGAC | GTGATCGGTA | AATACCACCC | GCACGGCGAC | TCCGCGGTAT | 200 |

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| ATGACACCAT | CGTTCGTATG | GCCCAGCCGT | TCTCGCTGCG | CTACATGCTG | 250 |
| GTGGATGGCC | AGGGGAACTT | CGGTTCAATC | GACGGCGACT | CCGCCGCGGC | 300 |
| AATGCGTTAT | ACGGAAATCC | GTCTGGCGAA | AATTGCTCAC | GAAGTGA | 347 |

2) INFORMATION FOR SEQ ID NO: 2279

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 362 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 43886

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2279

| | | | | | |
|------------|------------|------------|------------|------------|-----|
| AGAGCTCCTA | TCTGGATTAT | GCGATGTCGG | TCATTGTTGG | CCGTGCGCTG | 50 |
| CCAGATGTCC | GAGATGGCCT | GAAGCCGGTA | CACCGTCGCG | TACTTTACGC | 100 |
| CATGAACGTA | CTAGGCAATG | ACTGGAACAA | AGCCTATAAA | AAATCTGCCC | 150 |
| GTGTCGTTGG | TGACGTAATC | GGTAAATACC | ATCCCCATGG | TGACTCGGCG | 200 |
| GTCTATGACA | CGATCGTCCG | CATGGCGCAG | CCATTCTCGC | TGCGTTATAT | 250 |
| GCTGGTAGAC | GGTCAGGGTA | ACTTCGGTTC | TATCGACGGC | GACTCTGCGG | 300 |
| CGGCAATGCG | TTATACGGAA | ATCCGTCTGG | CGAAAATTGC | CCATGAACTG | 350 |
| ATGGCCGATC | TC | | | | 362 |

2) INFORMATION FOR SEQ ID NO: 2280

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 358 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 25922

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2280

| | | | | | |
|------------|------------|------------|-------------|------------|-----|
| CTCCTATCTG | GATTATGCGA | TGTCGGTCAT | TGTTGGCCGT | GCGCTGCCAG | 50 |
| ATGTCCGAGA | TGGCCTGAAG | CCGGTACACC | GTCGCGTACT | TTACGCCATG | 100 |
| AACGTACTAG | GCAATGACTG | GAACAAAGCC | TATAAAAAAT | CTGCCCGTGT | 150 |
| CGTTGGTGAC | GTAATCGGTA | AATACCATCC | CCATGGTGAC | TCGGCGGTTT | 200 |
| ATGACACGAT | CGTCCGTATG | GCGCAGCCAT | TCTCGCTGCG | TTACATGCTG | 250 |
| GTAGACGGTC | AGGGTAACTT | CGGTTCCATC | GACGGCGACT | CTGCGGCGGC | 300 |
| AATGCGTTAT | ACGGAAATCC | GTCTGGCGAA | AATTGCCCCAT | GAAGTGA | 350 |
| CCGATCTC | | | | | 358 |

2) INFORMATION FOR SEQ ID NO: 2281

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2281

CCCCCAGCTG GCGGGCGGTA TCGATGGGGG

30

2) INFORMATION FOR SEQ ID NO: 2282

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2282

AGRRGCIMAR ATGTATGA

18

2) INFORMATION FOR SEQ ID NO: 2283

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2283

ATITATGAYG GKITTCAGAG GC

22

2) INFORMATION FOR SEQ ID NO: 2284

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 bases
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2284

TCTGWGTRAC IGGYTCKGAG A

21

2) INFORMATION FOR SEQ ID NO: 2285

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2285

CMCCICCWGG TGGWGAWAC

19

2) INFORMATION FOR SEQ ID NO: 2286

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2286

AGTTGCTGTA TTAGGAAATG

20

2) INFORMATION FOR SEQ ID NO: 2287

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2287

TCGAAGTTGC TGTATTAGGA

20

2) INFORMATION FOR SEQ ID NO: 2288

1171

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1240 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*
- (B) STRAIN: BM4339

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2288

| | | | | | |
|-------------|------------|------------|------------|------------|------|
| TAGAAGCTGG | CTCGTTTTTT | TATAAATAAG | TTATTCGTTT | ATTTTTGTTT | 50 |
| TGTGCTAAAA | TATGAGAGTA | AATCACTGAA | CGATTTAGAA | TACAGGAGGA | 100 |
| CAATCTTTTG | AAGATTACTT | TACTATATGG | CGGACGCAGC | GCAGAGCAGA | 150 |
| GCATGAAGTG | TCCATTCTTT | CCGCATTTTC | AGTTTTAAAT | GCCATTTATT | 200 |
| ATAATTATTA | CCAAGTTCAA | CTCGTATTTA | TTACAAAAGA | AGGACAATGG | 250 |
| GTCAAAGGTC | CATTACTAAC | AGAAAAACCT | GCCAGCAAAG | ATGTCTTGCA | 300 |
| TCTTTCATGG | GACCCAAGTG | GACAGACAGA | GGAAGGCTTT | ACAGGAAAAG | 350 |
| TGATCAATCC | GGGCGAAATC | AAAGAAGAAG | GAGCCATCGT | TTTTCAGTT | 400 |
| TTACATGGGC | CAAACGGGGA | AGATGGAACG | ATCCAAGGCT | TCTTAGAGAC | 450 |
| ATTGAATATG | CCTTATGTCT | GCGCAGGCGT | ATTGACCAGT | GCATGTGCCA | 500 |
| TGGATAAAAT | CATGACCAAG | TATATTTTAC | AAGCTGCTGG | TGTGCCGCAA | 550 |
| GTTCCCTTATG | TACCACTACT | TAAGAATCAA | TGGAAAGAAA | ATCCTAAAAA | 600 |
| AGTATTTGAT | CAATGTGAAG | GTTCTTTGCT | TTATCCGATG | TTTGTCAAAC | 650 |
| CGGCGAATAT | GGGTCTAGT | GTCGGCATT | CAAAAGCAGA | AAACCGAGAA | 700 |
| GAGCTGCAAA | ATGCTTTAGC | AACAGCCTAT | CAGTATGATT | CTCGAGCAAT | 750 |
| CGTTGAACAA | GGAATTGAAG | CGCGCGAAAT | CGAAGTTGCT | GTATTAGGAA | 800 |
| ATGAAGACGT | TCGGACGACT | TTGCCTGGTG | AAGTCGTAAA | AGACGTAGCA | 850 |
| TTCTATGATT | ATGAAGCAAA | ATATATCAAT | AATAAAATCG | AAATGCAGAT | 900 |
| TCCAGCCGAA | GTGCCAGAAG | AAGTTTATCA | AAAAGCGCAA | GAGTACGCGA | 950 |
| AGTTAGCTTA | CACGATGTTA | GGTGGAAGCG | GATTGAGCCG | GTGCGATTTC | 1000 |
| TTTTTGTACAA | ATAAAAATGA | ATTATTCCTG | AATGAATTAA | ACTCTATGCC | 1050 |
| AGGATTTACG | GAGTTCAGTA | TGTACCCACT | CTTATGGGAA | AATATGGGCT | 1100 |
| TGAAATACGG | TGATTTGATT | GAAGAACTGA | TCCAGTTAGG | AATGAATCGA | 1150 |
| TACCATCAGC | GTCAATCTTT | TTTTGAAAAA | AATGAATAAA | GAGAAATAAA | 1200 |
| GAAGAGGCTG | GAGTGATTGC | GTAACCGCGT | TCATTCTAGC | | 1240 |

2) INFORMATION FOR SEQ ID NO: 2289

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2289

CACCGAAGAA GATGAAAAAA

20

2) INFORMATION FOR SEQ ID NO: 2290

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2290

TGGCACCGAA GAAGATGA

18

2) INFORMATION FOR SEQ ID NO: 2291

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2291

ATTTTGGCAC CGAAGAAGA

19

2) INFORMATION FOR SEQ ID NO: 2292

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2292

GAATCGGCAA GACAATATG

19

2) INFORMATION FOR SEQ ID NO: 2293

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1032 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*

(B) STRAIN: BM4147

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2293

| | | | | | |
|------------|------------|------------|------------|-------------|------|
| ATGAATAGAA | TAAAAGTTGC | AATACTGTTT | GGGGGTTGCT | CAGAGGAGCA | 50 |
| TGACGTATCG | GTAAATCTG | CAATAGAGAT | AGCCGCTAAC | ATTAATAAAG | 100 |
| AAAAATACGA | GCCGTTATAC | ATTGGAATTA | CGAAATCTGG | TGTATGGAAA | 150 |
| ATGTGCGAAA | AACCTTGCGC | GGAATGGGAA | AACGACAATT | GCTATTTCAGC | 200 |
| TGTACTCTCG | CCGGATAAAA | AAATGCACGG | ATTACTTGTT | AAAAAGAACC | 250 |
| ATGAATATGA | AATCAACCAT | GTTGATGTAG | CATTTTCAGC | TTTGCATGGC | 300 |
| AAGTCAGGTG | AAGATGGATC | CATACAAGGT | CTGTTTGAAT | TGTCCGGTAT | 350 |
| CCCTTTTGTA | GGCTGCGATA | TTCAAAGCTC | AGCAATTTGT | ATGGACAAAT | 400 |
| CGTTGACATA | CATCGTTGCG | AAAAATGCTG | GGATAGCTAC | TCCCGCCTTT | 450 |
| TGGGTTATTA | ATAAAGATGA | TAGGCCGGTG | GCAGCTACGT | TTACCTATCC | 500 |
| TGTTTTTGTT | AAGCCGGCGC | GTTCAGGCTC | ATCCTTCGGT | GTGAAAAAAG | 550 |
| TCAATAGCGC | GGACGAATTG | GACTACGCAA | TTGAATCGGC | AAGACAATAT | 600 |
| GACAGCAAAA | TCTTAATTGA | GCAGGCTGTT | TCGGGCTGTG | AGGTCGGTTG | 650 |
| TGCGGTATTG | GGAAACAGTG | CCGCGTTAGT | TGTTGGCGAG | GTGGACCAAA | 700 |
| TCAGGCTGCA | GTACGGAATC | TTTCGTATTC | ATCAGGAAGT | CGAGCCGGAA | 750 |
| AAAGGCTCTG | AAAACGCAGT | TATAACCGTT | CCCGCAGACC | TTTCAGCAGA | 800 |
| GGAGCGAGGA | CGGATACAGG | AAACGGCAAA | AAAAATATAT | AAAGCGCTCG | 850 |
| GCTGTAGAGG | TCTAGCCCGT | GTGGATATGT | TTTTACAAGA | TAACGGCCGC | 900 |
| ATTGTACTGA | ACGAAGTCAA | TACTCTGCCC | GGTTTCACGT | CATACAGTCG | 950 |
| TTATCCCCGT | ATGATGGCCG | CTGCAGGTAT | TGCACTTCCC | GAAGTGATTG | 1000 |
| ACCGCTTGAT | CGTATTAGCG | TTAAAGGGGT | GA | | 1032 |

2) INFORMATION FOR SEQ ID NO: 2294

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2294

AAACGAGGAT GATTGATTG

20

2) INFORMATION FOR SEQ ID NO: 2295

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2295

TTGAGCAAGC GATTTCGG

18

2) INFORMATION FOR SEQ ID NO: 2296

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1029 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*
- (B) STRAIN: V583

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2296

| | | | | | |
|-------------|------------|------------|-------------|------------|------|
| ATGAATAAAA | TAAAAGTCGC | AATTATCTTC | GGCGGTTGCT | CGGAGGAACA | 50 |
| TGATGTGTCTG | GTAAAATCCG | CAATAGAAAT | TGCTGCGAAC | ATTAATACTG | 100 |
| AAAAATTCTGA | TCCGCACTAC | ATCGGAATTA | CAAAAAACGG | CGTATGGAAG | 150 |
| CTATGCAAGA | AGCCATGTAC | GGAATGGGAA | GCCGATAGTC | TCCCCGCCAT | 200 |
| ATTCTCCCCG | GATAGGAAAA | CGCATGGTCT | GCTTGTCATG | AAAGAAAGAG | 250 |
| AATACGAAAC | TCGGCGTATT | GACGTGGCTT | TCCCCGGTTT | GCATGGCAAA | 300 |
| TGCGGGGAGG | ATGGTGCGAT | ACAGGGTCTG | TTTGAATTGT | CTGGTATCCC | 350 |
| CTATGTAGGC | TGCGATATTC | AAAGCTCCGC | AGCTTGCAATG | GACAAATCAC | 400 |
| TGGCCTACAT | TCTTACAAAA | AATGCGGGCA | TCGCCGTCCC | CGAATTTCAT | 450 |
| ATGATTGAAA | AAGGTGACAA | ACCGGAGGCG | AGGACGCTTA | CCTACCTGT | 500 |
| CTTTGTGAAG | CCGGCACGGT | CAGGTTCGTC | CTTTGGCGTA | ACCAAAGTAA | 550 |
| ACAGTACGGA | AGAACTAAAC | GCTGCGATAG | AAGCAGCAGG | ACAATATGAT | 600 |
| GGAAAAATCT | TAATTGAGCA | AGCGATTTCG | GGCTGTGAGG | TCGGCTGCGC | 650 |
| GGTCATGGGA | AACGAGGATG | ATTTGATTGT | CGGCGAAGTG | GATCAAATCC | 700 |
| GGTTGAGCCA | CGGTATCTTC | CGCATCCATC | AGGAAAACGA | GCCGGAAAAA | 750 |
| GGCTCAGAGA | ATGCGATGAT | TATCGTTCCA | GCAGACATTC | CGGTGAGGA | 800 |
| ACGAAATCGG | GTGCAAGAAA | CGGCAAAGAA | AGTATATCGG | GTGCTTGAT | 850 |
| GCAGAGGGCT | TGCTCGTGTT | GATCTTTTTT | TGCAGGAGGA | TGGCGGCATC | 900 |
| GTTCTAAACG | AGGTCAATAC | CCTGCCCCGT | TTTACATCGT | ACAGCCGCTA | 950 |
| TCCACGCATG | GCGGCTGCCG | CAGGAATCAC | GCTTCCCGCA | CTAATTGACA | 1000 |
| GCCTGATTAC | ATTGGCGATA | GAGAGGTGA | | | 1029 |

2) INFORMATION FOR SEQ ID NO: 2297

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2297

TTCAGGAGGG GGATCGC

17